é

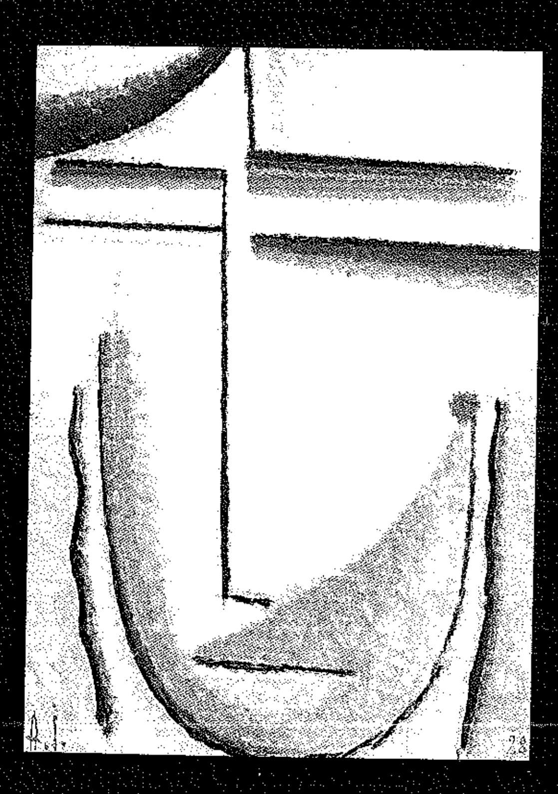
propósito general de esta obra es el de elaborar un
xto que signifique un aporte en el campo de la
ca de la metodología, partiendo de manera articulada
el tratamiento de aspectos teórico-filosóficos como
os. La perspectiva original de este libro es la de
ir a una estrategia de reconstrucción socio histórica,
ente utilizada en la enseñanza de los métodos de las
s sociales. A partir de esta reconstrucción histórica,
ca elaborar una propuesta que no se circunscriba a un
onjunto especializado de saberes entre todos aquellos
n la actualidad, constituyen el patrimonio
ológico de las ciencias sociales.
decir que hemos buscado presentar el panorama
sta de los métodos y las técnicas de investigación,
ndo especial atención a la actualización del
miento, intentando incluir en este sentido a los más
es desarrollos metodológicos y la utilización de
mentos tecnológicos.
parado pedagógicamente para cursos de Metodología
Ciencias Sociales en carreras de grado, es también
ndado para cursos de pos grados y para todos
s que necesitan de las herramientas metodológicas
reas de investigación.

ISBN 978-950-04-2868-2
9 789500 428682

METODOLOGÍA DE LAS CIENCIAS SOCIALES

Alberto Marradi
Nélida Archenti
Juan Ignacio Piovani

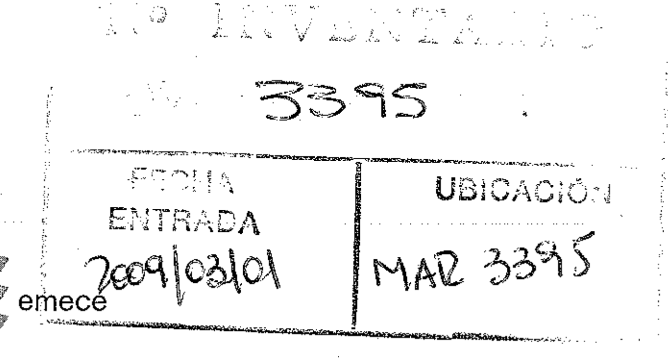
METODOLOGÍA DE LAS CIENCIAS SOCIALES



emecé

Marradi, Archenti y Piovani

Metodología de las Ciencias Sociales



emecé

Marradi, Alberto
Metodología de las ciencias sociales - Alberto Marradi, Nélida Archenti, Juan Ignacio Piovani – 1ª ed.– Buenos Aires : Emecé Editores, 2007.
328 p. ; 16x25 cm.

ISBN 978-950-04-2868-2

1. Metodología de la investigación 2. Ciencias Sociales I. Nélida Archenti, Juan Ignacio Piovani II. Título
CDD 001.42

Independencia 1668, C 1100 ABQ, Buenos Aires, Argentina
www.editorialplaneta.com.ar

Diseño de cubierta: *Departamento de Arte de Editorial Planeta*
1ª edición: abril de 2007
Impreso en Printing Books,
Mario Bravo 835, Avellaneda,
en el mes de marzo de 2007.

IMPRESO EN LA ARGENTINA / PRINTED IN ARGENTINA
Queda hecho el depósito que previene la ley 11.723
ISBN: 978-950-04-2868-2

ÍNDICE GENERAL

Nota de autores

Alberto Marradi

Profesor titular ordinario de Metodología de las Ciencias Sociales y director de la carrera de Sociología de la Università di Firenze, Italia. En el marco de la Asociación Internacional de Sociología, fue presidente del Research Committee in Conceptual and Terminological Analysis y vicepresidente del Research Committee in Methodology. Actualmente es presidente del Comité de Coordinación de los doctorados italianos en Sociología y director de la Maestría en Metodología de la Investigación Social, Università di Bologna —representación en Buenos Aires— y Universidad Nacional de Tres de Febrero. Asimismo, dirige el Sistema de Escuelas de Excelencia de la Asociación Italiana de Sociología. Ha publicado *Concetti e metodi in scienza politica*, Florencia, Giuntina (1980); *L'analisi monovariata*, Milán, Angeli (1992); *Linee guida per l'analisi bivariata dei dati nelle scienze sociali*, Milán, Angeli (1994); *Raccontar storie*, Roma, Bonanno (2005), y *Metodologia delle scienze sociali*, Bolonia, Il Mulino (2007). Además, es coautor de 3 volúmenes y compilador de 5 colecciones de ensayos metodológicos, y autor de docenas de ensayos en varios idiomas.

Nélida Archenti

Licenciada en Sociología (Universidad de Buenos Aires) y posgraduada en Sociología y Metodología de la Investigación Social (Fundación Bariloche, Argentina). Profesora titular regular en la carrera de Ciencia Política, Facultad de Ciencias Sociales, Universidad de Buenos Aires. Profesora de Metodología de la Investigación Social en diversos posgrados de Argentina y del exterior. Investigadora del Instituto Gino Germani, dependiente de la Facultad de Ciencias Sociales de la Universidad de Buenos Aires. Ha publicado *Actualidad del pensamiento sociopolítico clásico* (con Luis Aznar), Buenos Aires, EUDEBA (4ª edición, 2004), y *Situación de la mujer en la sociedad argentina: formas de organización en Capital Federal*, Buenos Aires, Fundación Friedrich Naumann (1987). Asimismo, es coautora de *Igualdad de derechos, igualdad de oportuni-*

dades, igualdad de ejercicio, Santiago de Chile, Fundación Instituto de la Mujer (2004) y autora de una gran variedad de artículos y capítulos de libros en los que se tratan temas de su especialidad.

Juan Ignacio Piovani

Doctor en Metodología de las Ciencias Sociales (Università di Roma "La Sapienza") y Magíster en Métodos Avanzados de Investigación Social y Estadística (The City University, London). Profesor titular ordinario en la Facultad de Humanidades de la Universidad Nacional de La Plata. Profesor titular en varios posgrados de la Argentina y del exterior. Investigador del Centro Interdisciplinario de Metodología de las Ciencias Sociales (CIMeCS), Facultad de Humanidades, Universidad Nacional de La Plata. Coordinador de la Maestría en Metodología de la Investigación Social, Università di Bologna —representación en Buenos Aires— y Universidad Nacional de Tres de Febrero. Ha publicado *Alle origini della statistica moderna. La scuola inglese di fine Ottocento,* Milán, Angeli (2006) y coeditado *El sudeste asiático: una visión contemporánea,* Buenos Aires, Eduntref (2004). Además, es autor de varios artículos y capítulos de libros sobre temas de su especialidad.

Prefacio

Hacia fines de 2004, cuando comenzamos a trabajar en el proyecto que dio lugar a este libro, nos planteamos en primera instancia la necesidad de hacer explícitos sus objetivos y de definir una serie de ideas que hicieran las veces de principios rectores que dotaran a nuestra propuesta de una identidad particular. El propósito general fue elaborar un texto que significara, en múltiples sentidos, un aporte en el campo de la didáctica de la metodología. Ahora que el libro se ha materializado, creemos fundamental dar a conocer al lector este conjunto de principios subyacentes que informaron nuestro trabajo. A esta tarea dedicamos los párrafos siguientes.

La primera cuestión relevante surgió de la reflexión acerca del concepto mismo de metodología. Resulta oportuno declarar que compartimos aquellas posturas que asignan a la metodología un espacio central en un *continuum* de análisis crítico ubicado entre el estudio de los postulados epistemológicos que hacen posible el conocimiento social y la elaboración de las técnicas de investigación. En este sentido, ocuparse de metodología es encontrarse en una continua tensión dialéctica entre los polos de este *continuum*: si se obtura la dimensión epistemológica, la metodología se reduce a un conjunto de prácticas no controladas intelectualmente, y si se abandona el aspecto técnico, se transforma en una especulación abstracta sobre las ciencias sociales, incapaz de incidir sobre las actividades de investigación concretas. Es por eso que nuestro texto incluye, de manera articulada, el tratamiento de aspectos tanto teórico-filosóficos (epistemológicos, ontológicos y gnoseológicos) como técnicos, y tiende puentes entre ambos dominios.

Una revisión crítica de los textos disponibles en el campo de la metodología muestra que muchos de ellos, en cambio, se limitan al tratamiento de los argumentos epistemológicos o bien al de las cuestiones técnicas. Estas tendencias "unidireccionales" suelen darse en circunstancias en que los términos más elementales del campo —'metodología', 'método', 'técnica'— no se definen con precisión o incluso se toman como intercambiables. No deberá sorprender al lector, por lo tanto, ver que hemos dedicado un capítulo completo a clarificarlos y a reflexionar acerca de estas importantes (y fundacionales) decisiones conceptuales y terminológicas.

Por otra parte, en el tratamiento de los temas tanto filosóficos como técnicos hemos recurrido a una estrategia de reconstrucción socio-histórica, raramente utilizada en la enseñanza de los métodos de las ciencias sociales. En cierto sentido, esta desatención por las cuestiones históricas resulta comprensible, teniendo en cuenta que en general, y en el marco de significativas limitaciones de tiempo, se deben abordar en los cursos de grado y posgrado una serie de temas técnicos orientados al diseño y la realización de investigaciones empíricas. Sin embargo, la importancia de conocer dichos temas en perspectiva histórica no es una cuestión marginal; además de sensibilizar al estudiante acerca de su carácter de "construcciones sociales", facilita un uso más crítico y creativo de los instrumentos de la investigación. Esto, en parte, porque ayuda a comprender los procesos de producción y reproducción de sentidos en torno de los conceptos y herramientas metodológicas.

A partir de tal reconstrucción histórica hemos buscado, además, elaborar una propuesta que no se circunscribiera a un solo conjunto especializado de saberes entre todos aquellos que en la actualidad constituyen el patrimonio metodológico de las ciencias sociales. A diferencia de muchos textos especializados —por cierto valiosos— que se han difundido en los últimos años, y que se concentran en uno u otro de los tipos de investigación que habitualmente se conocen como cuantitativo y cualitativo, hemos optado por presentar un panorama pluralista de las estrategias que se siguen en la producción de conocimiento en las ciencias sociales. Intentamos, entendiendo críticamente su lenguaje y sus respectivos instrumentos conceptuales y operativos, ir más allá de la tradicional distinción de los métodos apenas citada. A cada uno de ellos, que hemos preferido llamar "estándar" (cuantitativo) y "no-estándar" (cualitativo), le hemos dedicado un mismo nivel de atención, y ambos han sido abordados con una particular aproximación crítica. Tal aproximación se funda en la idea según la cual todo proceso de investigación involucra un conjunto de decisiones no lineales y de distintos niveles de complejidad, que, a su vez, ponen en juego saberes impersonales (técnicos, objetivados) y saberes personales (tácitos, no explicitados y difícilmente explicitables).

Cuando se trata, por ejemplo, de los métodos estándar, dicho enfoque crítico implica, entre otras cosas, una denuncia de sus usos dogmáticos y ritualizados, así como un reconocimiento del rol central del investigador en la selección de las herramientas más adecuadas y en la interpretación significativa de los resultados de su empleo.

En el caso de los métodos no-estándar, la actitud crítica, reconociendo el lugar clave que el investigador tiene en todo el proceso de indagación, hace notar simplemente que la apelación a la creatividad y a la subjetividad como recursos genuinos de investigación no puede tomarse como excusa legitimadora de un "todo vale" metodológico. También se destaca que en estos tipos de investigación tiene presencia —aunque probablemente en una medida relativa mucho menor que en la aproximación estándar— el conocimiento de tipo técnico.

Por otra parte, sostenemos que la sistematicidad en el trabajo y el compromiso por hacer transparentes las decisiones que han hecho posible cierto tipo de conclusiones en el marco de una investigación deberían ser característicos de cualquier tipo de estudio, sea estándar o no. A propósito de hacer transparentes las decisiones de los investigadores, creemos oportuno aclarar que, si bien nos inscribimos en el marco de una perspectiva de denuncia de las desigualdades de género en el campo científico, y a pesar de ser conscientes de que tales desigualdades se cristalizan en el lenguaje, hemos decidido no recurrir a fórmulas del estilo "el/la investigador/a" sino utilizar términos masculinos en sentido genérico para referirnos tanto a hombres como a mujeres, con el único fin de hacer menos engorrosa la lectura del texto.

Resta decir que hemos buscado presentar el panorama pluralista de los métodos y las técnicas de investigación al que hemos hecho referencia prestando especial atención a la actualización del conocimiento, intentando incluir, en este sentido, los más recientes desarrollos metodológicos y la utilización de instrumentos tecnológicos.

Finalmente, cabe indicar —para cerrar este breve prefacio— que el libro se estructura en torno a 4 módulos analíticamente diferenciables. Si bien en todos ellos se trata una variedad de temas articulados: cuestiones históricas, fundamentos teórico-filosóficos, instrumentos conceptuales y operativos, etcétera, cada uno se caracteriza por el énfasis dado a algún aspecto distintivo. Así, el primer módulo, que incluye los capítulos 1 a 4, se centra en los fundamentos teóricos y filosóficos de la investigación social, los debates metodológicos contemporáneos y la clarificación de términos clave. El segundo —del capítulo 5 al 9 inclusive— comprende el desarrollo de los principales instrumentos conceptuales de la investigación social. En el tercer módulo (capítulos 10 a 14) se abordan las más difundidas técnicas de recolección de información, y el cuarto (capítulos 15 a 17) se dedica fundamentalmente a los problemas de análisis.

Los autores
Buenos Aires, enero de 2007.

CAPÍTULO 1

TRES APROXIMACIONES A LA CIENCIA

1.1. Galileo y el método experimental

Muchos filósofos e historiadores opinan que la ciencia moderna empieza a fines del siglo XVI. En efecto, en esta fecha Galileo asesta el golpe final a la visión aristotélica de la ciencia, ya criticada por Ockham y otros filósofos empiristas ingleses. Según esta visión —compartida por muchos filósofos helenistas y árabes y por toda la Escolástica— la tarea de la ciencia era identificar la naturaleza de cada especie de objeto del conocimiento, separando lo que es esencial y fijo (la sustancia) de lo que es accidental y variable.

Por el contrario, para Galileo —y para la visión que se convierte en estándar en los tres siglos siguientes—, la tarea de la ciencia es formular, controlar y decidir aserciones acerca de relaciones entre las propiedades de los objetos. El científico tiene que hallar la forma matemática que estas relaciones tienen en la naturaleza (Galileo, remontándose a Pitágoras a través de Platón, afirma que "el libro de la naturaleza está escrito en caracteres matemáticos").

Ya que las relaciones matemáticas entre las propiedades no son directamente accesibles a la observación, el científico tiene que saber interrogar hábilmente la naturaleza y penetrar sus secretos, eliminando todos los accidentes que podrían afectar la nitidez de la relación matemática: para ver con qué velocidad los cuerpos caen al suelo se debe imaginar un vacío sin resistencia del aire; para ver con qué velocidad se mueven los cuerpos a lo largo de un plano inclinado se deben imaginar cuerpos perfectamente esféricos y ausencia de rozamiento, y así análogamente.

Estos experimentos de Galileo son mentales, porque él nunca logró un vacío absoluto o ausencia de rozamiento. Pero él sentó las bases del MÉTODO EXPERIMENTAL que desarrollaron sus seguidores, en particular Torricelli.

Galileo puede también ser considerado uno de los padres de la aproximación nomotética[1] a la ciencia, en cuanto las relaciones matemáticas que buscaba

[1] Según la cual la tarea de los científicos es encontrar —o formular— leyes de la naturaleza. Véase más adelante.

entre una pareja de propiedades dada valían, obviamente, en cualquier tiempo y lugar.

En su forma ideal, un experimento se obtiene sólo cuando se observen los efectos de variaciones controladas de una propiedad llamada 'operativa' en otra propiedad llamada 'experimental', mientras se mantienen constantes todas las otras propiedades que podrían influir sobre la segunda. Es decir, se ponen en práctica medidas que impidan variaciones en los estados de las propiedades de este grupo durante todo el experimento.[2] Este modelo realiza lo que Galileo llamaba "sabia interrogación" de la naturaleza. La sabiduría consiste no sólo en elegir la propiedad operativa más apropiada —un problema que compete a la teoría— sino también en producir cambios perfectamente controlados en ésta, y en bloquear todas las variaciones en las otras propiedades que se sabe o se supone puedan influir sobre la experimental, manteniéndolas efectivamente constantes —un problema exquisitamente técnico.

Al realizar un experimento de esta forma ideal, el científico podrá averiguar la forma pura (es decir, sin influencias de otra fuente) de la relación entre la propiedad operativa y la experimental, y también determinar la dirección de esta relación: la operativa tiene una influencia causal en la otra. A partir de un experimento (ideal) se cree poder afirmar, por ejemplo: "El aumento de la propiedad A determina una reducción de la propiedad B según la siguiente una función matemática: ..."

Los objetos de muchas ciencias físicas poseen una característica que no tienen los objetos de las ciencias sociales: son FUNGIBLES (intercambiables). Es decir, se puede dar por sentado que los objetos del mismo tipo siempre reaccionen de la misma manera bajo las mismas condiciones. Esta característica, que no pertenece al método experimental sino a la naturaleza del dominio estudiado por la física, la química y otras ciencias parecidas, permite sin embargo extrapolar el resultado de un experimento más allá del caso específico, generalizando este resultado a todos los objetos considerados del mismo tipo en la disciplina dentro de la cual el experimento ha sido diseñado.

Como veremos, esto conlleva una consecuencia estructural extremadamente relevante. En el experimento ideal se opera sobre un único objeto, considerado típico, o representativo de todos los objetos de su especie. Se pueden registrar todas las informaciones relevantes (para establecer una relación matemática) en dos vectores:[3] en uno se ponen los estados en la propiedad operativa en una secuencia de puntos de tiempo durante el experimento, en el otro los estados en la experimental en los mismos puntos del tiempo. Se grafican estos datos en una abscisa y una ordenada —herramientas intelectuales que Descartes proporcionó a la naciente comunidad científica— y se traza una

[2] Las mayores dificultades que se encuentran al investigar mediante experimentos en las ciencias físicas son las de hallar y aplicar estas medidas de neutralización.

[3] Un vector es una secuencia ordenada de informaciones relativas al mismo referente. Eso tiene un patrón que expresa la forma básica de la predicación: su referente es el objeto del discurso (de quién, o de qué, se habla) y en sus posiciones se registran sistemáticamente las informaciones acerca de los atributos de este objeto.

curva o una recta representando la relación matemática. El análisis matemático de los datos de un experimento se desarrolla típicamente con parejas de vectores de este tipo.[4]

1.2. Límites del modelo experimental

Este modelo experimental, que contribuyó poderosamente al desarrollo de algunas ciencias físicas, tiene sin embargo algunos límites.

Un límite teórico consiste en el hecho de que jamás se puede excluir con seguridad que otras propiedades, no incluidas en el modelo, influyan en la propiedad experimental. En otras palabras, siendo infinitas las propiedades existentes en la naturaleza, al diseñar cualquier experimento la gran mayoría de ellas se deben considerar irrelevantes,[5] basándose en las teorías existentes, es decir, se dejan variar sin intervención, no se las hace variar, pero tampoco se impide su variación. Debido a la presencia de estas innumerables propiedades que varían libremente (y que eventualmente podrían ejercer influencia), el modelo no se puede jamás dar por "cerrado".

Los límites prácticos consisten en que no siempre es técnicamente posible controlar a la perfección las variaciones en la propiedad operativa, y todavía menos posible es neutralizar perfectamente la influencia de las propiedades que se deberían mantener efectivamente constantes. Además, no siempre se puede excluir que la experimental reaccione influyendo en la operativa, de manera que la relación que se da por unidireccional sea en cierta medida bidireccional.

En resumidas cuentas, el método experimental es particularmente adecuado para investigar relaciones monocausales, típicamente diacrónicas. Por el mismo motivo, no es adecuado para investigar relaciones funcionales, típicamente sincrónicas. Tampoco es adecuado para investigar sistemas complejos de relaciones causales, cuando hay una tupida red de interacciones entre muchas propiedades que ejercen influencias diferentes en diferentes combinaciones: en estas situaciones, fijar la influencia que ejerce una simple propiedad en una función matemática precisa, bloqueando las variaciones de las otras, es a menudo imposible, y en todo caso artificial. Aún más inadecuado es el modelo

[4] Naturalmente, éste es el modelo más simple, cuando hay una sola propiedad operativa.

[5] Todas las propiedades de todos los objetos involucrados en un experimento (piénsese en las herramientas) son un conjunto infinito, y sólo se considera un pequeño subconjunto de ellas para declararlo explícitamente irrelevante. Todas las otras propiedades no son consideradas ni siquiera para declararlas irrelevantes. Esto no excluye que algunas podrían ser en realidad relevantes: nunca se puede excluir con certeza que una propiedad cualquiera ejerza una influencia, siquiera indirecta. En efecto, en muchas ciencias de la naturaleza el afianzado edificio teórico —y también el conocimiento personal y tácito de los investigadores— permite considerar con bastante tranquilidad como irrelevantes la mayor parte de las infinitas propiedades que pueden concebirse. Pero los avances teóricos cambian continua e imperceptiblemente el marco, y lo que parecía irrelevante se vuelve relevante para la generación siguiente. Finalmente, este proceso garantiza el continuo progreso de las ciencias físicas.

experimental para las relaciones de tipo teleológico, cuando las propiedades se refieren a sujetos que tienen planes, objetivos y voluntad, y que típicamente no reaccionan de manera uniforme a un tratamiento experimental.

Estos límites del método experimental, que no le impiden ser muy provechosos en varias ciencias físicas, tienen en cambio graves consecuencias en las ciencias sociales. La razón es que no se puede soñar con aislar una pareja de propiedades bloqueando todas las otras, porque ellas seguirán influyendo en la experimental, y/o en la operativa,[6] y/o en la relación entre ellas. Además, muchas propiedades importantes en las ciencias sociales no pueden ser modificadas *ad libitum* por el experimentador: o son fijas —como la raza— o varían según un patrón en que no se puede intervenir —como la edad.

Desconociendo esos serios límites, algunos han tratado de aplicar el método experimental a problemas de las ciencias sociales, introduciendo modificaciones y adaptaciones que a veces han resultado sutiles e ingeniosas.[7] Pero incluso los resultados más interesantes no pueden generalizarse más allá de cada experimento particular, debido a la artificialidad de las situaciones experimentales y al número muy limitado de sujetos, carentes de cualquier representatividad, a los que se aplica el experimento. Por éstos y por los otros límites señalados arriba, los pretendidos "experimentos" en las ciencias sociales distan mucho de la verdadera naturaleza del método experimental: lo demuestra el hecho de que nadie en la comunidad científica, ni siquiera el experimentador, cree seriamente que un experimento cualquiera haya identificado la forma matemática de la relación entre dos propiedades de individuos o de grupos.

Además, incluso si se creyese que en aquella situación se había logrado aislar la relación entre A y B de toda influencia extraña, este resultado sería relativo a ella y no extensible a todos los casos de la relación entre A y B en cualquier tiempo y lugar. Esto se debe a que en las ciencias sociales no se puede pensar seriamente que los objetos sean fungibles, y que se pueda estudiar un individuo o una sociedad extrapolando los resultados a cualquier otro individuo o a toda otra sociedad.

1.3. La matriz de datos y el supuesto atomista

Por lo tanto, al tiempo de la revolución científica de Galileo faltaba una herramienta intelectual que cumpliese para las ciencias sociales la misma función que la pareja de vectores (dos secuencias paralelas de datos acerca de los estados de un objeto en dos propiedades) llevaba a cabo en el experimento.

[6] En las ciencias sociales se habla de 'independiente' en lugar de 'operativa' y de 'dependiente' en lugar de 'experimental' (véase apartado 15.1).

[7] Véanse las obras clásicas de Campbell y Stanley (1963), Cook y Campbell (1979) y el capítulo 4 de Corbetta (2003), que concluye su detallada y articulada descripción de diseños experimentales destacando algunos límites que también se señalan en el texto: artificialidad del ambiente y falta de representatividad (2003: 151-4). Una crítica más amplia en Pawson y Tilley (1997).

Además, en los tiempos de Galileo —y a lo largo de los dos siglos siguientes— en las ciencias sociales faltaban unidades de medida u otras maneras de cuantificar los estados, mientras que para las propiedades físicas esenciales (el espacio, el tiempo, el peso) se tenían UNIDADES DE MEDIDA desde hacía siglos. En rigor de verdad, hasta el siglo XIX, estas unidades eran en su mayoría locales, no universales,[8] pero bastaban para establecer una relación matemática que no era difícil de traducir en las unidades locales de cada comunidad científica.

La herramienta básica del enfoque que sigue siendo el estándar en las ciencias sociales fue desarrollada a mediados del siglo XVIII gracias a un profesor de la universidad de Göttingen en Alemania —Alfred Achenwall—, quien proporcionaba a su soberano informaciones acerca de recursos económicos y militares de otros estados alemanes,[9] poniendo reinos, ducados y ciudades libres en las filas y sus estados en varias propiedades en las columnas de una rudimentaria matriz.

La matriz es una herramienta, elaborada por matemáticos en el siglo XVI, formada por un cruce entre un haz de vectores paralelos[10] horizontales y un haz de vectores paralelos verticales. Más allá de su uso matemático para sistematizar y acelerar los cálculos, tiene muchos otros usos en varias disciplinas.[11]

Su importancia para las ciencias sociales consiste en el hecho de que permite tratar la variabilidad entre objetos del mismo tipo (seres humanos, instituciones, otros productos de la cultura) sin renunciar a formas de análisis estadístico. Como un vector simple permite dar forma sistemática a la predicación (atribución de estados) acerca de un solo objeto, así la matriz —y más precisamente un tipo de ésta, llamado "matriz de datos", que veremos en el apartado 6.1—, poniendo diferentes referentes del mismo tipo en sus filas y la misma secuencia de propiedades en sus columnas, permite relacionar rápida y sistemáticamente un sinnúmero de estados de los mismos objetos/sujetos en distintas propiedades, controlando así si existe una relación cualquiera entre estas propiedades.

Una vez que se ha registrado en una matriz (veremos cómo en el capítulo 7), los estados de n sujetos en p propiedades, ésta nos permite establecer cuáles son las relaciones existentes entre cualquier pareja de propiedades,[12] aplicando la técnica estadística adecuada a la naturaleza de las propiedades, y con validez restringida a los sujetos estudiados.

[8] La universalización de las medidas físicas empezó con el establecimiento del metro patrón en un congreso en París durante los primeros años de la Revolución Francesa. A partir de allí, a lo largo del siglo XIX, se estableció una verdadera red de unidades de medida estrechamente interrelacionadas para una gran cantidad de propiedades estudiadas por la física.

[9] Hasta la proclamación del imperio en 1871, Alemania estaba dividida en tres reinos (Prusia, Baviera, Sajonia) y muchos Estados menores.

[10] En un haz de vectores paralelos, los referentes son diferentes pero del mismo tipo, y las informaciones están rigurosamente en la misma secuencia.

[11] Por ejemplo, una tabla bivariable es a su vez una matriz.

[12] Con técnicas oportunas y adoptando algunos supuestos, se pueden cuantificar relaciones ternarias y todavía más complejas.

Esta perspectiva, que permitió desarrollar técnicas de análisis de datos de gran sofisticación en el siglo XX, descansa sobre un sistema de supuestos tan conectados entre sí que pueden ser designados con un rótulo global (SUPUESTO ATOMISTA). Analíticamente, ellos aparecen en la lista siguiente:

- cada caso (sea un individuo humano, un estado, o un ejemplar de cualquier otro tipo de unidad) que se ponga como referente en una fila de la matriz se puede descomponer perfectamente en sus estados en las propiedades registradas en la matriz;
- cada estado en una propiedad, una vez transformado en dato en la matriz, es perfectamente separable de su detentador;
- cada estado en una propiedad es totalmente independiente de los otros estados de su detentador en todas las otras propiedades (estén o no registradas en la matriz);
- cada estado al cual, aplicando una definición operativa,[13] le fue asignado un valor simbólico (código) transformándolo en un dato de la matriz, es tratado como perfectamente igual a cada otro estado que haya recibido el mismo valor (numérico). En otras palabras, todos los datos con el mismo valor (en la misma variable)[14] son perfectamente asimilables entre ellos.

Sólo este supuesto atomista hace posible el análisis de datos en la forma que lo conocemos, porque permite construir distribuciones de frecuencia de los valores de una variable y establecer relaciones de asociación entre dos o más variables. Cada tabla y cada diagrama pueden ser tratados —en la fase del análisis— como un universo cerrado, sin hacer referencia ni al resto de la matriz ni a los objetos (siendo, en la mayoría de los casos, sujetos) detentadores de los estados.

1.4. El método de la asociación

Naturalmente, en el siglo XVIII, del cual estamos hablando, las técnicas de análisis de datos eran extremadamente elementales: piénsese que la mediana fue imaginada por Wundt y por Galton al final del siglo siguiente y que el promedio, para ahorrar computaciones manuales, era a menudo sustituido por el *midrange* (es decir la mitad de la suma del máximo y del mínimo valor de la distribución). Además, si incluso en las ciencias físicas las propiedades que se medían eran todavía pocas, y las unidades eran a menudo —como se vio— locales, no había ninguna unidad de medida específica en las ciencias sociales. Las únicas propiedades cuantitativas a disposición eran los resultados de conteos

[13] Una definición operativa es un conjunto de reglas cuya aplicación permite transformar un estado en la realidad en un dato en la matriz. Véase apartado 7.1.

[14] Una variable es un conjunto de signos (datos) que representan en la matriz los estados correspondientes en la propiedad: véase capítulo 7.

(número de hogares, número de suicidas, y demás). Sólo los economistas podían agregar (a los conteos del número de unidades compradas/vendidas de un cierto producto) todas las propiedades expresadas en dinero (costos, precios, inversiones, y demás) y por lo tanto muy temprano —a fines del siglo XVIII— estuvieron listos para intentar imitar el modelo de ciencia galileana ya prevaleciente en las ciencias físicas.

En un cierto sentido, ellos estaban incluso adelantados en relación con las ciencias físicas, porque —como se señaló en la nota 8— sólo en el siglo XIX los físicos se dotaron de unidades de medida universalmente aceptadas para todas las propiedades consideradas relevantes en ese momento.

A mediados de ese siglo empezaron también los intentos de medir propiedades de seres humanos, a medio camino entre la naturaleza física y la psique —por ejemplo, las reacciones a sonidos, colores, olores. Las razones por las que estas tentativas no podían alcanzar algo parecido a la medición en sentido estricto serán profundizadas en el apartado 8.4. Sin embargo, estos primeros ingeniosos intentos de medir propiedades no físicas revisten una importancia que va más allá de su escaso éxito.

A partir de los años veinte del siglo XX, un psicólogo que venía de la física a través de la psicofísica, Louis Leon Thurstone, trató de aplicar los mismos procedimientos (véase apartado 8.4.) para establecer unidades de medida de las actitudes. Imaginó varias técnicas sofisticadas, y proclamó en 1928 que "las actitudes pueden ser medidas".[15] En realidad, a pesar de ser ingeniosas y complejas, estas técnicas están muy alejadas de la impersonalidad y repetibilidad que son el requisito de un acto de medición (véase apartado 8.4).

Pero Thurstone había abierto el camino, y pronto otras propuestas más simples y eficaces serían planteadas por Likert, Guttman y otros (véase apartado 8.5). No se había conseguido la deseada medición, pero antes de finalizar el siglo XX la psicología —y por imitación otras ciencias sociales— gozaba de un abanico de técnicas para cuantificar, al menos en forma aproximada, sus propiedades más específicas, como valores, actitudes, opiniones, transformándolas en variables cuantitativas.

A pesar del hecho de que en el siglo XIX se empezaban a usar tablas bivariables —una herramienta adecuada para analizar relaciones entre variables no cuantitativas— y que al final del siglo el estadístico escocés George Udny Yule hubiese abordado el problema de cuantificar este tipo de relaciones con coeficientes adecuados, la mayoría de los investigadores veían y ven en la disponibilidad de variables cuantitativas —o supuestas y tratadas como tales— y de herramientas adecuadas para analizarlas el rasgo que caracterizaba el enfoque que en el siglo XX devino estándar en las ciencias sociales, y que por eso fue y es denominado 'cuantitativo', contraponiéndolo[16] al que se llama 'cualitativo' —del cual hablaremos pronto.

[15] El título de un célebre artículo suyo.

[16] Como distinción relativa al método con el cual se procede en la investigación, la contraposición cuantitativo/cualitativo parece superficial y no apropiada. En una matriz de datos, se en-

Considerando que:

- todas las formas de recolección de informaciones en el así llamado enfoque estándar en las ciencias sociales transforman —mediante una definición operativa y su ejecución— estas informaciones en datos en una matriz;
- esta transformación y codificación, a pesar de que inevitablemente tergiversa la realidad (véase capítulo 7), consigue una ventaja medular: la posibilidad de establecer relaciones entre propiedades fundándolas en algo más que impresiones, y cuantificando la fuerza de estas relaciones;
- a diferencia del experimento, no se puede establecer de manera impersonal (es decir, sin intervención del conocimiento humano) *la dirección* de una eventual influencia, por lo tanto las técnicas de análisis permiten establecer la presencia de asociaciones[17] entre variables pero no de relaciones causales;

se podría llamar MÉTODO DE LA ASOCIACIÓN este que sustituye al método experimental en las ciencias sociales siguiendo en el marco del enfoque estándar (o "concepción heredada", o "consenso ortodoxo") de la ciencia.

1.5. El enfoque que se contrapone a la visión estándar

Se puede rastrear la aproximación que ahora se llama "cualitativa" hasta el pensamiento del filósofo italiano Vico (1708), la perspectiva hermenéutica, los historicistas alemanes Dilthey, Windelband y Rickert, Max Weber (el primero en destacar que cada fenómeno social no tiene una explicación única), la fenomenología de Schutz, el interaccionismo simbólico de Thomas, Mead y Blumer, la etnometodología de Garfinkel: sobre estos autores y enfoques se detendrá el apartado 2.3.[18]

cuentran variables "cualitativas" junto a variables "cuantitativas". El instrumento de recolección por excelencia de los investigadores "cuantitativos" es el cuestionario "cerrado". Pero no existe cuestionario que se limite a plantear preguntas sobre aspectos cuantitativos o cuantificables de la realidad. La gran mayoría de las preguntas que se encuentran en los cuestionarios típicos conciernen a aspectos *no* cuantificables, o sea "cualitativos", de la realidad. Por otro lado, los investigadores de la escuela más prestigiosa de la sociología "cualitativa", la escuela de Chicago, no despreciaron aspectos cuantitativos: por ejemplo, para evaluar el nivel de acatamiento a las leyes en el período prohibicionista, se instruía a los jóvenes investigadores a *contar* el número de botellas vacías de cerveza y otras bebidas alcohólicas que se veían en los tachos de basura que se encuentran cerca de la escalera de seguridad en la mayoría de los condominios norteamericanos.

[17] El término 'asociación' fue introducido por Yule (1900) con un significado más amplio del término 'correlación', que Galton y Pearson reservaban a relaciones entre variables cuantitativas. 'Asociación' es el género: las especies, según la naturaleza de las variables involucradas, toman el nombre de correlación, cograduación, concordancia (véase capítulo 15). Sobre la querella entre Yule y Pearson, véase Piovani (2006).

[18] Sobre el enfoque hermenéutico véase también Lulo (2002); sobre el pensamiento de Schutz véase también Soldano (2002).

Mientras que el método experimental y el método que proponemos llamar "de la asociación" son fuertemente estructurados, en el sentido de que presentan estrechas conexiones entre los supuestos, los conceptos, los términos, las técnicas y el tipo de aserciones que se producen, en la aproximación comúnmente llamada "cualitativa" conviven formas de investigación muy diferenciadas: el único rasgo seguramente común a todas es un rasgo negativo: la no adopción de —que en la mayoría de los casos es un rechazo[19] a adoptar— los supuestos fundamentales de la visión estándar de la ciencia (la que se desarrolló en las ciencias físicas a partir de Galileo: véase arriba). Muchos investigadores no aceptan pagar los precios semánticos de representaciones demasiado simples, o irrealistas, de los fenómenos humanos y sociales[20] para gozar de las ventajas sintácticas de la matriz de datos y de las herramientas de análisis conectados con ella.[21]

Más allá del rechazo, esos investigadores a menudo comparten un valor en positivo: la sensibilidad hacia el aporte cognoscitivo y los derechos humanos de los sujetos conocidos —una sensibilidad que a veces se constituye en verdadera "epistemología del sujeto conocido" (Vasilachis 2000). De esta orientación, que no dudaría en llamar ética, se siguen algunos de los caracteres comunes que —a pesar de la gran variedad de la cual se hablaba— se pueden identificar en la mayoría de los caminos NO ESTÁNDAR de investigación. A continuación, se intentará esta identificación, y también se pondrán dichos caracteres en un orden de sucesión que tenga algún sentido.

Lo que nos parece el punto de partida es la

a) orientación para reducir al mínimo la separación entre ciencia y vida cotidiana, por lo tanto, la fuerte propensión a abandonar las torres de marfil y tomar contacto directo con los sujetos mientras llevan a cabo sus actividades habituales, dejándolos expresar sus visiones del mundo en sus propios términos y poniéndolos en el centro de los reportes de investigación, una gran atención por problemas de significado y una actitud de escucha frente a la realidad en vez de una rígida limitación a modelos *a priori*.

Esta tendencia a seguir de cerca la vida cotidiana conlleva una

b) fuerte dependencia del contexto: toda investigación queda estrechamente confinada a la situación específica que investiga.

[19] No todos los investigadores perciben esto como un rechazo (término que implica una connotación valorativa e incluso emotiva): para algunos, solamente se trata de una renuncia que límites objetivos hacen inevitable.

[20] Se da por sentado que —mientras el "método de la asociación" se usa también en las ciencias físicas, cuando no se puede usar el método experimental— el enfoque del cual estamos hablando pertenece únicamente a (algunos) cultores de ciencias humanas y sociales.

[21] Una elección plenamente legítima si no se transforma en un rechazo *a priori* de ciertas herramientas cualquiera sean los objetos de investigación y los intereses del investigador —rechazo que tal vez cubre la ignorancia y esconde el deseo de ahorrar la fatiga del aprendizaje.

Esto a su vez conlleva una

c) preferencia por los problemas micro —que en algunos investigadores no excluye la ambición de enfrentar sobre esta base también los problemas macro [véanse por ejemplo Geertz (1983) y algunos ensayos recogidos en Knorr-Cetina y Cicourel (1981)];
d) orientación marcadamente idiográfica, descriptiva —que en muchos autores no implica renunciar a formular conceptos de alcance y pretensiones generales;
e) orientación marcadamente inductiva:[22] en vez de "verificar hipótesis" preformuladas, el investigador abre de par en par los ojos sobre las escenas y tiende las orejas a todos los mensajes que la situación le presenta.

La fuerte "contextualización" de la investigación conlleva además:

f) la preferencia por la comprensión global de "provincias delimitadas de significado" [la expresión es de Schutz (1945)] y de situaciones específicas más que por la institución de relaciones causales lineales entre variables —y aún menos por la "verificación" de aserciones singulares en torno de asociaciones entre estas variables;
g) el hecho de que la causación, si es contemplada, es considerada un proceso que se reconstruye por medio de narraciones y no por medio de relaciones estadísticas entre instantáneas escalonadas en el tiempo o entre variables relevadas en el mismo momento (como en la familia de la asociación: véase capítulo 15).

La falta de procedimientos sólidamente establecidos y generalmente aceptados —de un "método científico" codificado y ritualizado al cual atenerse— conlleva:

h) una gran importancia de las cualidades y capacidades del investigador y de sus colaboradores, que en la investigación del tipo estándar son a menudo considerados meros ejecutores. La experiencia en investigaciones previas, el conocimiento del contexto que se estudia, la flexibilidad, la creatividad, la intuición, son mucho más importantes que las (eventuales) herramientas técnicas para determinar la calidad de la investigación;
i) la dificultad para desarrollar este género de investigación sin caer en lo banal, en lo gratuito, en lo anecdótico, y sin "volverse nativo" (*go native*: algunos investigadores adoptan tan profundamente el punto de vista de los sujetos estudiados que se vuelven o se sienten miembros de su comunidad / de su grupo / de su "provincia delimitada de significado" —lo que puede ser una excelente solución a problemas existenciales individuales, pero es casi siempre una pérdida para la comunidad científica). Todos estos peligros son advertidos también por muchos exponentes de esta orientación.

[22] Esta orientación, entiéndase, es compartida por la mayoría de los investigadores del enfoque "no estándar" —en particular por los que siguen la *grounded theory* (véase Glaser y Strauss 1967)— pero no por todos.

Naturalmente, al construir esta secuencia de caracteres comunes pusimos de relieve las conexiones entre ellos. En la práctica de las investigaciones de orientación no estándar, y en la autorrepresentación de los que siguen estos caminos, algunas conexiones entre caracteres son más tenues de lo que aparecen en la reconstrucción, y aceptarla integralmente podría ser problemático para algunos. Como se dijo, no existe —y parece improbable que se desarrolle— un patrimonio de conceptos, técnicas y prácticas comúnmente reconocidos y empleados.

CAPÍTULO 2

LOS DEBATES METODOLÓGICOS CONTEMPORÁNEOS

2.1. Introducción

Durante el siglo XX el debate metodológico se centró fundamentalmente en los conceptos de CUALIDAD y CANTIDAD. El foco puesto en la relación controversial de este par conceptual hace a este debate tributario de las discusiones en la filosofía y la metodología de la teoría social del siglo XIX (Marx, Durkheim, Weber).

Uno de los principales ejes de la controversia se estructuró alrededor del abordaje elegido para dar cuenta de la relación entre individuo y sociedad, dando lugar, por un lado, a una perspectiva centrada en el sujeto y, por otro, a una basada en la estructura social. La primera tendió a poner el acento en la razón y la acción del sujeto, orientando las investigaciones hacia la búsqueda de la comprensión de las motivaciones y las decisiones individuales, considerando a los individuos capaces de la construcción y la interpretación de las conductas generadas en la interacción social. La segunda tendió a explicar la acción individual a partir de la estructura, por la pertenencia e integración a un todo social basado en normas.

El sustrato de estos enfrentamientos es de tipo ontológico, es decir, remite a alguna concepción subyacente de la realidad o, como sostiene Schwandt (1994: 132), a "supuestos acerca de cómo debe ser el mundo para que lo podamos conocer". Sus raíces se encuentran en la sociología clásica europea: desde una perspectiva durkheimiana, se trata de un mundo social regulado por normas con un orden semejante al natural; mientras que en el abordaje weberiano, se trata de un mundo caótico que los sujetos organizan para poder conocerlo. Para algunos hay un orden preexistente cognoscible, para otros el orden social es una construcción humana.

La diversidad de estos enfoques y sus desacuerdos dieron lugar a debates epistemológicos sobre el estatus científico de las ciencias sociales, y metodológicos sobre los modos de producir y validar el conocimiento científico en estas disciplinas.

Al calor de esta controversia comenzaron a definirse, durante el siglo XX, dos "bandos" comprometidos en un "combate de religiones" —tal como lo in-

dica figuradamente Marradi (1997b). Por un lado, uno aglutinado bajo la aceptación y promoción de la medición como "mejor" medio para asegurar la cientificidad de las prácticas de investigación de las ciencias sociales; por el otro, uno comprometido con el rechazo radical de cualquier intento de cuantificación de la realidad social (véase apartado 1.5).

Ferrarotti (1983: 9) llamaba irónicamente "sociografía"[1] a las prácticas y a la perspectiva sociológica del primer grupo, a cuyos exponentes juzgaba "responsables de [la] degradación [de la sociología]", por haberla condenado a la "pérdida de su conciencia problemática" y haberla hecho "funcional a los intereses económicos dominantes y a la lógica del mercado". Las prácticas del segundo grupo eran, en cambio, habitualmente denunciadas como no científicas, y sus defensores, como representantes de una *nouvelle vague* antimetodológica (Statera 1984) que proponía un "edén" imposible de alcanzar (Leonardi 1991).

Estos debates cobraron actualidad en la década de 1960, en circunstancias que Giddens (1979) califica como "disolución del consenso ortodoxo" de las ciencias sociales, cuando terminaron por popularizarse algunas antinomias que en cierta medida aún nos acompañan: cuantitativo versus cualitativo, explicación versus comprensión, objetividad versus subjetividad, neutralidad versus participación; en definitiva, descripción de la sociedad "tal cual es" versus crítica y transformación de la sociedad actual.

A partir de los años ochenta esta controversia epistemológica comenzó a perder fuerza. En las ciencias sociales empezó a considerarse que la cuestión de los MÉTODOS CUANTITATIVOS y CUALITATIVOS no se resolvía en el plano de las discusiones filosóficas sobre la realidad, sino en el plano de la racionalidad de medios afines entre: un problema cognitivo de interés, un diseño de investigación apropiado al problema y los instrumentos técnicos más adecuados para resolverlo. En este contexto se fue imponiendo lentamente lo que Bryman (1988) denomina "argumento técnico":[2] los métodos cuantitativos y cualitativos son apropiados para alcanzar distintos objetivos cognitivos y tratar problemas de índole diferente, y la tarea del investigador no es apegarse acríticamente a un modelo, sino tomar las decisiones técnicas pertinentes en función del problema de investigación que enfrenta (Marradi 1992). También es posible imaginar problemas de investigación cuyo abordaje requiera de una combinación de métodos, lo que se conoce habitualmente como TRIANGULACIÓN METODOLÓGICA (véase apartado 2.4).

[1] El término "sociografía" (y su equivalente en otras lenguas) ha sido usado con diferentes sentidos. En las ciencias sociales uno de sus primeros usos se atribuye al pensador escocés Patrik Geddes, quien llamaba así a su particular enfoque, que él mismo describía como una síntesis de sociología y geografía. Como se advierte en la cita, el sentido de Ferrarotti es más bien despectivo: busca poner en evidencia la pobreza de ciertas perspectivas sociológicas que se reducen al uso de estadísticas y con fines meramente descriptivos. Por otra parte, ésta es la definición del *Webster Encyclopedic Unabridged Dictionary* (1996): rama de la sociología que usa datos estadísticos para describir fenómenos sociales.

[2] Además de los argumentos epistemológico y técnico aquí referidos, en la literatura metodológica también se encuentran argumentos ontológicos, axiológicos y gnoseológicos como fundamento de la distinción cualitativo/cuantitativo (Piovani *et al.* 2006).

En la actualidad no son pocos los autores que creen que una clasificación de los métodos basada en el criterio de cualidad/cantidad resulta poco útil y debería abandonarse (a propósito de esto, véase por ejemplo nuestra propuesta en el capítulo 1). Sin embargo, los usos terminológicos no son fácilmente modificables, y se sigue apelando muy frecuentemente a los rótulos de "estudio cualitativo" o "estudio cuantitativo".

Por otra parte, es de hacer notar que las tradiciones teóricas y epistemológicas a las que se ha recurrido usualmente para dar cuenta de los enfoques cuantitativo y cualitativo —positivismo e interpretativismo respectivamente— son menos monolíticas y más complejas de lo que lo que se suele reconocer. Con el fin de aportar a una comprensión más profunda de las raíces y desarrollo histórico de los debates epistemológicos y metodológicos contemporáneos de las ciencias sociales, presentamos a continuación una breve reconstrucción de dichas tradiciones, enfatizando sus matices y tensiones internas. Partimos del análisis de las perspectivas canónicas (representadas en este caso por el positivismo y sus sucedáneos) para luego pasar al conjunto de enfoques críticos que se han postulado como alternativas epistemológico-metodológicas.

2.2. El debate intrapositivista

La perspectiva generalista encontró su versión metodológica más acabada en el positivismo, siguiendo el camino marcado por Auguste Comte, Gabriel Tarde y Herbert Spencer, y más aún por Emile Durkheim —heredero del pensamiento de Bacon y Descartes. Durkheim (1895), a la luz del modelo de las ciencias naturales, estableció una analogía entre el objeto de las ciencias sociales —el hecho social— y las cosas, y predicó la necesidad de tratar metodológicamente al primero igual que a las segundas. Esta necesidad anclaba en el convencimiento de la existencia de un único modelo científico válido para todas las ciencias, perspectiva que se conoce como MONISMO METODOLÓGICO (véase apartado 3.2). La propuesta de tomar como modelo a la física suponía que la realidad social también estaba regida por leyes universales, susceptibles de ser descubiertas con la aplicación del mismo método científico.

En el marco de este paradigma fue tomando forma una estrategia metodológica habitualmente conocida como cuantitativa, pero más adecuadamente definible en este caso como método de la asociación (véanse apartados 1.3 y 1.4). Éste se caracteriza por el recurso a un conjunto de instrumentos conceptuales y operativos para la investigación empírica que permitirían —en principio— cumplir objetivos análogos a los que en la física cumplía el experimento —asumiendo que éste no podía ser aplicado al nuevo objeto. Dichos objetivos, desde el punto de vista cognitivo, constituyen el fin principal de la ciencia moderna en clave galileana (véase apartado 1.1): en definitiva, se pretendía cuantificar aspectos de la realidad social para luego poder establecer sistemáticamente relaciones entre ellos, con el objetivo final de postular leyes generales sobre el funcionamiento de la sociedad y de los fenómenos sociales.

En el ideario tardo-positivista, la ley científica estaba desprovista de cualquier contenido metafísico que le diera un tinte de necesidad inherente; era más bien la generalización de una secuencia dada de fenómenos empírica y repetidamente observados, y dotada por lo tanto de regularidad. Por otra parte, tal secuencia —en tanto rutina perceptiva— era la base empírica de la explicación causal. Al respecto, cabe hacer notar que para los máximos mentores del instrumental técnico de la investigación social cuantitativa (de la asociación) de fines del siglo XIX —especialmente Karl Pearson— la idea de contingencia era más adecuada que la de causalidad. Para ellos, la causalidad no era un principio dicotómico, de suma cero (un hecho es causa o no de otro), sino una gradación de distintos niveles posibles de relación entre los fenómenos. Desde su perspectiva, la tarea del científico era justamente la de determinar las formas y grados de la relación entre los fenómenos estudiados (a través de la correlación), y generalizar los resultados a partir de una lógica inductiva (Piovani 2006).

A la idea de GENERALIZACIÓN se sumaban las de OBJETIVIDAD y EXTERNALIDAD: el carácter externo y autónomo de la realidad exigía la objetividad como requisito para alcanzar conocimiento válido. Así, la idea de la neutralidad valorativa se impuso como una de las características elementales del conocer científico.

El carácter empírico de la actividad científica, basada en la medición y el manejo de los datos como sustento de la explicación, se fue afianzando en los ciencias sociales durante el siglo XX, y adquirió carácter predominante —particularmente en EE.UU.— a través de la difusión de la técnica del sondeo (*survey*). El desarrollo de la estadística y la aplicación de la teoría de la probabilidad a las técnicas de muestreo, que permitía predecir con importante aproximación la conducta de grandes poblaciones a partir de muestras relativamente pequeñas, contribuyó a la fascinación por el número y la medida.

En los años cuarenta, y bajo el liderazgo intelectual de Robert Merton y Paul Lazarsfeld, se desarrolló en la Universidad de Columbia el denominado *survey research*, iniciando una línea de investigación basada en sondeos que contribuyó con importantes aportes a la teoría sociopolítica.[3] Por otra parte, el *survey research* fue el exponente más típico de la operativización, en las ciencias sociales, de lo que se conoce como visión estándar de la ciencia —*standard view* (véanse apartados 1.3 y 1.4). La visión estándar, heredera de la tradición positivista, ejerció un predominio indiscutido en la epistemología de mediados del siglo XX, especialmente en el mundo anglosajón (Mulkay 1979, Outhwaite 1987, Piovani 2002). Si bien no se trata de una posición carente de matices, se puede encuadrar genéricamente en lo que se conoce como neopositivismo y falsacionismo, y ligar a la obra de Carnap (1939), Hempel (1966), Nagel (1961) y muy especialmente Popper (1934, 1963).

[3] Con posterioridad, la aparición de la tecnología informática permitió manejar una cantidad muy grande de datos en la búsqueda de relaciones multivariables por medio de programas estadísticos, en particular de datos provenientes de encuestas y censos.

A principios del siglo XX era común sostener que el conocimiento científico difería de otras descripciones y explicaciones del mundo porque se derivaba de los hechos: a partir de las repetidas observaciones de la realidad, siguiendo una mecánica inductiva, se podían alcanzar generalizaciones sobre los fenómenos estudiados. En cambio, Popper va a proponer una alternativa que se conoce habitualmente como falsacionismo.

Como él mismo señala en *Conjectures and Refutations* (1963), en la Viena de la primera posguerra comenzó a interesarse por el siguiente problema: "*¿cuándo una teoría debe ser considerada científica?* o *¿existe un criterio del carácter o estatus científico de una teoría?*" Para entender el sentido de esta preocupación resulta importante indagar acerca del porqué de este interés en encontrar un CRITERIO DE DEMARCACIÓN que distinguiera al conocimiento científico. A diferencia de sus predecesores positivistas y de sus contemporáneos neopositivistas del Círculo de Viena, él no se interesaba tanto en la distinción entre ciencia y metafísica; su objetivo primordial era el de diferenciar la ciencia de la pseudociencia. Para Popper (1963) la pseudociencia es el conocimiento que se presenta como científico, pero que en realidad tiene más que ver con los mitos primitivos.[4] En su perspectiva, el conocimiento no comienza por una observación ateórica cuya repetición permite una generalización inductiva, sino por la postulación de una conjetura o hipótesis que pretende describir o explicar algún aspecto de la realidad y a partir de la cual se derivan —deductivamente— enunciados observacionales que permitirán su contrastación empírica. Por lo tanto, la ciencia sigue un camino HIPOTÉTICO-DEDUCTIVO. La característica singular del conocimiento científico es que esta conjetura debe ser pasible de FALSACIÓN, es decir, debe poder ser demostrada falsa (y nunca definitivamente verdadera) a través de enunciados observacionales específicos. Según Popper (1972), además, es a través de los mecanismos de falsación que la ciencia progresa.

La consagración del método hipotético-deductivo como única vía válida para la producción de conocimiento científico revitalizó la idea del monismo metodológico: para Popper la ciencia se ocupa de explicaciones conformadas por sistemas de hipótesis que han resistido las pruebas de falsación, y estas pruebas sólo pueden hacerse a través de un único y mismo método.

El impacto de las ideas de Popper en las ciencias sociales fue importante: ellas fueron tomadas como base epistemológica por parte de la sociología académica norteamericana de la época y contribuyeron a dar sustento a lo que Giddens (1979) denomina "consenso ortodoxo" de las ciencias sociales, dominante especialmente en Estados Unidos hacia mediados del siglo XX, y articulado en torno del estructural-funcionalismo de Talcott Parsons y otros. Por otra parte, la afinidad de las ideas de Popper con estas corrientes sociológicas no se limitaba a las esferas epistemológicas y metodológicas. Como dice Aron (1996: 244), en el ámbito de lo social y lo político Popper buscaba "defender e

[4] Según Popper (1963), ejemplos de grandes teorías pseudocientíficas de la época eran la teoría marxista de la historia, la teoría psicoanalítica de Freud y la psicología individual de Adler.

ilustrar con una argumentación lógica [...] una actitud reformista contra la actitud revolucionaria", talante profundamente compartido por los sociólogos de la época enrolados en las tradiciones teóricas del consenso.

Las propuestas de Popper pasaron a formar parte —acríticamente— de los manuales de metodología de la investigación social que se difundieron a escala global, en los que se las presentaba como "el" método científico (véase apartado 3.2). Esto trajo aparejadas algunas consecuencias negativas para la investigación social empírica: además de desacreditar cualquier práctica que se alejara del método hipotético-deductivo, significó el desplazamiento del problema de investigación del centro de la escena investigativa y la entronización de las hipótesis —muchas veces banales, postuladas de forma forzada y ritualista— como rectoras del proceso de investigación.

A pesar de su influencia, casi hegemónica hasta principios de los años sesenta y todavía vigente en muchos sentidos, y aun reconociendo la importancia de su crítica al inductivismo, la visión popperiana de la ciencia —y del progreso científico— no es inmune a las críticas. Una de las más sistemáticas y reconocidas se encuentra en la obra de Thomas Kuhn, particularmente en *The Structure of Scientific Revolutions* (1962).

Kuhn entiende que la ciencia debe ser estudiada y concebida como un proceso histórico; no se trata de un producto lógico-racional que se produce en el vacío, sino de un conjunto de prácticas sociales —históricamente condicionadas— que tienen lugar en el seno de una comunidad científica. Su concepto clave es el de PARADIGMA. Según la posdata de 1969 a su libro apenas citado, paradigma puede entenderse en un doble sentido: a) a nivel más general, como un conjunto de generalizaciones simbólicas, modelos heurísticos, valores comunes y soluciones ejemplares compartidas por una comunidad científica en un momento dado; b) en un sentido restringido, como decisión ejemplar relativa a la solución exitosa de algún tipo específico de problema científico.

A partir de su enfoque histórico, Kuhn distingue entre una etapa preparadigmática (en la que no hay acuerdos generales sobre el objeto de estudio, los científicos se encuentran relativamente aislados entre sí, y proliferan las corrientes que brindan criterios alternativos de investigación e interpretación, todos en pugna por prevalecer) y una etapa paradigmática (cuando un paradigma logra imponerse —generando amplios consensos en la comunidad científica y dando lugar en consecuencia a un período denominado de CIENCIA NORMAL). Pero pueden surgir momentos de crisis —producto de anomalías— que implican la puesta en cuestión de algunos de los consensos básicos del paradigma, no problematizados en las etapas de ciencia normal. Estas crisis pueden retrotraer el estado de la ciencia a una situación preparadigmática, y eventualmente desembocar en una REVOLUCIÓN CIENTÍFICA, por medio de la cual logrará afianzarse un nuevo paradigma, en cierto sentido inconmensurable con los anteriores.

En línea con la noción positivista tradicional relativa a un cierto grado de retraso madurativo de las ciencias sociales con respecto a las naturales, Kuhn y otros autores han planteado que, a diferencia de las físicas, por ejemplo, las ciencias sociales nunca han pasado de la etapa preparadigmática. La controver-

sia cualitativo/cuantitativo, para ellos, no sería más que un síntoma de la falta de consensos paradigmáticos sobre las formas de investigar y validar el conocimiento científico. Por otra parte, la idea de paradigmas en pugna —como se ha planteado en los primeros párrafos de este capítulo— ha constituido uno de los argumentos más utilizados (argumento epistemológico, como ya se ha visto) para sostener la distinción entre ambos tipos de métodos.

La imagen de la ciencia como resultado de prácticas sociales históricamente situadas —aspecto saliente del planteo de Kuhn— generó importantes debates al interior de la visión estándar. Probablemente, el último gran intento por preservar algunas de sus ideas básicas, ante las consecuencias que la obra de Kuhn traía, se encuentre en los escritos del epistemólogo húngaro Imre Lakátos. En un cierto sentido, su pensamiento refleja una apuesta por conciliar el falsacionismo popperiano con la perspectiva sociohistórica de Kuhn: "la unidad de análisis de Lakátos es lo que llama un 'programa de investigación', noción que tiene a la vez componentes sociológicos y lógicos, y que parece haberse originado en una conjunción de aspectos kuhnianos y popperianos" (Klimovsky 1994: 373).

Los programas de investigación incluyen un núcleo duro (conjunto de hipótesis fundamentales que la comunidad científica ha decidido no poner en cuestión) y un cinturón protector (constituido por una serie de hipótesis auxiliares que a través de su adaptación o modificación permiten la adecuación entre el núcleo duro y las observaciones). Además de esto, otro aspecto en el que se hace evidente el intento de Lakátos por dar respuesta a Kuhn, salvando algunos elementos del pensamiento de Popper, es el relativo al desarrollo histórico de la ciencia. Para esto recurre como estrategia a la distinción entre una historia interna (que comprende el estudio de las cuestiones lógicas y metodológicas por medio de las cuales se produce y valida el conocimiento científico) y una historia externa (que atañe a factores aparentemente extracientíficos, como la ideología, la cultura, el desarrollo económico, etcétera). Sin embargo, para explicar la evolución de la ciencia en tanto empresa racional, Lakátos —más cercano a la visión canónica— establece una jerarquía en la que otorga preeminencia a la historia interna.

A pesar del esfuerzo de Lakátos por salvar (algo de) la epistemología convencional a través de su falsacionismo sofisticado, la "revolución" iniciada con la publicación de la obra de Kuhn (1962) siguió desmoronando de manera ineluctable el edificio epistemológico y metodológico construido sobre las bases del positivismo/neopositivismo/falsacionismo. Además de nuevos embates como los de Feyerabend (1970), expresado por medio de su propuesta de una teoría anarquista del conocimiento, una de las consecuencias más relevantes se dio en términos del desarrollo de perspectivas históricas y sociológicas para el estudio de la ciencia. En el marco de ellas florecieron las miradas constructivistas y relativistas. Las críticas al positivismo, por mucho tiempo confinadas a posiciones minoritarias —y casi todas en el campo de las ciencias sociales y las humanidades— y de alcance limitado (por ejemplo la hermenéutica, que será analizada en la próxima sección), se generalizaron y alcanzaron los mismos cimientos (o núcleo duro) del proyecto positivista. Revitalizadas por la obra de

Kuhn, que tuvo un gran impacto en las ciencias sociales, las epistemologías propias de estas disciplinas también creyeron tener algo para decir acerca de la "impoluta" historia interna de la ciencia, incluso rechazando la distinción misma entre historia interna y externa como estrategia para dar cuenta de la ciencia y de su desarrollo histórico.

2.3. Las críticas al positivismo y los enfoques no estándar

Como se sugirió en párrafos precedentes, las críticas al positivismo no surgieron con la obra de Kuhn. Éstas tienen en cambio una larga y rica tradición; sólo que tal como hemos apuntado, su influencia en la epistemología y la metodología fue relativamente marginal hasta décadas recientes: ella se limitó en general al mundo de las ciencias sociales y las humanidades, e incluso en muchos casos —por ejemplo, la hermenéutica— no buscó rechazar al positivismo *per se*, sino contener su intromisión en disciplinas para cuyos objetos se lo consideraba inadecuado.

El término 'HERMENÉUTICA' deriva del griego ερμηνευμα, que significa interpretación. Si bien sus orígenes pueden rastrearse hasta los estudios literarios de los retóricos de la antigua Grecia y las exégesis bíblicas de la Patrística, la hermenéutica —en su sentido moderno— hace referencia a una interpretación profunda que involucra una relación compleja entre sujeto interpretador y objeto interpretado.

El primero en explorar la importancia de la interpretación más allá de la exégesis de los textos sagrados fue Schleiermacher, durante la primera mitad del siglo XIX. Él inauguró una tradición que luego continuarían los historicistas Dilthey, Windelband y Rickert, entre fines del siglo XIX y principios del XX, y más recientemente Gadamer, Apel y Ricoeur, entre otros.

La tradición hermeneútica adquirió relevancia metodológica y epistemológica para las ciencias sociales en la medida que destacó la especificidad de su objeto de estudio, y la consecuente necesidad de métodos propios para abordarlo, ligándolos al problema de la interpretación. En este sentido se destaca la figura de Dilthey, que extendió el dominio de la hermenéutica a todos los fenómenos de tipo histórico y promovió la comprensión (*verstehen*), por oposición a la explicación (*erklärung*),[5] como un movimiento desde las manifestaciones exteriores de la conducta humana hacia la exploración de su significado intrínseco.

La especificidad del objeto y la necesidad de métodos propios fue una idea heredada del filósofo de la historia italiano Giambattista Vico, quien había rechazado la posibilidad de aplicar el método cartesiano a los fenómenos humanos (o cívicos, en su propio lenguaje). Al absolutismo y objetivismo metodológico cartesiano también se opuso Immanuel Kant (1781), quien consideraba que los sentidos no constituían el único medio de la percepción.

[5] La distinción entre el explicar, como característico de las ciencias naturales; el conocer (*erkennen*), como típico de la filosofía y la teología, y el comprender, como propio de las ciencias históricas, se debe a Droysen (1868), quien ejerció una importante influencia en Dilthey.

Para Vico, en el conocimiento del orden humano de la realidad es fundamental la sabiduría práctica (o frónesis, del griego φρονησις), que ya Aristóteles había señalado como la aplicación del buen juicio a la conducta humana, en oposición a la sabiduría (o sofía, σοφια), que hacía referencia al conocimiento de las causas, o de por qué las cosas son como son. Por otra parte, Vico destacaba que este tipo de fenómenos estaba gobernado por aspectos en cierto sentido imponderables, en cuya investigación no se podía seguir un esquema lineal, predefinido, tal como había postulado Descartes.

Se va a ir consolidando así una distinción entre el conocimiento del mundo material (la naturaleza), para el cual los hermeneutas no negaban los principios y métodos generalizadores de corte positivista, y el conocimiento de los fenómenos espirituales (humanos), que requerían de una interpretación profunda y que por sus características y complejidad no podían reducirse a leyes universales. Esta distinción llegó al clímax en el marco de la Escuela de Baden, cuando Windelband primero, y Rickert después, propusieron la distinción entre aproximaciones NOMOTÉTICAS (típicas de las ciencias de la naturaleza), que tienen por objeto la postulación de leyes generales basadas en procesos causales uniformes, y aproximaciones IDIOGRÁFICAS (típicas de las ciencias humanas o del espíritu, *geisteswissenschaften*), cuyo objeto es el estudio de fenómenos cambiantes que deben ser interpretados en su especificidad y por lo tanto situados contextualmente.

Entre fines del siglo XIX y principios del siglo XX, esta tradición intelectual maduró en las ciencias sociales empíricas de la mano del sociólogo alemán Max Weber. Así como el pensamiento de Durkheim se puede señalar como el más importante antecedente específicamente sociológico del enfoque estándar, la obra de Max Weber, vista en retrospectiva, constituye el antecedente más destacado de la perspectiva no estándar.

En su obra *Wirtschaft und Gesellschaft*, publicada póstumamente en 1922, Weber define a la sociología como una ciencia que pretende entender, interpretándola, la acción social. Pero dado que para él la acción es un comportamiento subjetivamente significativo, promueve la COMPRENSIÓN como el método adecuado para investigarla. Influido por su maestro Rickert, intenta distinguir el abordaje sociológico del modelo físico, destacando el carácter histórico de los fenómenos sociales. Asimismo, se opone al objetivismo imperante en las ciencias de la naturaleza; en el ámbito de la cultura hechos y valores se entrecruzan, y esto se refleja incluso en la investigación, cuando el estudioso selecciona un tema de interés.

Desde un punto de vista más instrumental y operativo, los denominados métodos cualitativos de investigación, enmarcados en la corriente interpretativa de raíz weberiana, abrevan del método etnográfico de la antropología clásica. En efecto, si bien se toma en ocasiones a los trabajos de Tocqueville y Le Play como predecesores de la investigación cualitativa —por oposición a los estudios cuantitativos de Quetelet— lo habitual es considerar como sus antecedentes más directos las prácticas etnográficas de la antropología clásica (y de la es-

cuela sociológica de Chicago). En la antropología de fines del siglo XIX y principios del siglo XX ya se recurría a formas de trabajo de campo que muchos consideran —al menos superficialmente— análogas a ciertas prácticas de la actual investigación social no estándar.[6]

La antropología clásica se enmarcaba dentro de corrientes colonialistas; sus estudios etnográficos eran fundamentalmente de tipo descriptivo, con la mirada puesta en otras culturas en las que el investigador —"etnógrafo solitario"— se trasladaba para recolectar datos orientados al estudio de un otro cultural "primitivo" que, organizados en notas de campo, eran luego volcados en informes de investigación llamados habitualmente "monografías".[7]

Contemporáneamente al desarrollo de la antropología clásica, durante las primeras décadas del siglo XX, florece en Estados Unidos la sociología de la ESCUELA DE CHICAGO. En este ámbito, y a partir de una confrontación entre el uso de los métodos estadísticos y el estudio de caso, el sociólogo polaco Florian Znaniecki propone la distinción metodológica entre inducción enumerativa (en la que los casos son tratados como ejemplos de colectivos) e inducción analítica (en la que cada caso ilumina aspectos de una teoría general). La investigación de Thomas y Znaniecki (1918-1920), *The Polish Peasant in Europe and America*, marcará un hito del método etnográfico: la inducción analítica pasará a ser considerada como intrínsecamente etnográfica pues, aunque en los estudios también se utilizaran métodos cuantitativos, éstos tenían un estatus subordinado frente a los métodos cualitativos.

En el Departamento de Sociología de la Universidad de Chicago (que por mucho tiempo fue Departamento de Sociología y Antropología) también desarrollaron sus trabajos Park y Burgess, quienes destacaron el carácter idiográfico de los estudios sociales y dieron lugar a la sociología urbana, y Mead y Blumer, fundadores del INTERACCIONISMO SIMBÓLICO.

[6] Sin embargo, se discute si tal enfoque puede considerarse genuinamente cualitativo en el sentido contemporáneo (véase apartado 10.4). Vidich y Lyman (2000), por ejemplo, han señalado que la investigación de la antropología clásica y de la sociología de la Escuela de Chicago nació sin la intención de "comprender al otro". Por otra parte, la escuela funcionalista inglesa (exponente saliente de la antropología clásica), en la que se enmarcan figuras como Malinowski, Radcliffe-Brown y Evans Pritchard, estuvo decisivamente influenciada por pensadores positivistas de la talla de Comte, Spencer y Durkheim. En la Escuela de Chicago, por su parte, y según Platt (1985), toda la evidencia indica que las propuestas comprensivistas weberianas eran completamente desconocidas y no fueron tomadas como sustento teórico-metodológico de la investigación empírica que allí se desarrollaba.

[7] La antropología transita luego por perspectivas evolucionistas —basadas en teorías del desarrollo pensado en etapas sucesivas, y en la idea de un otro "subdesarrollado" como objeto de estudio— hasta que, avanzado el siglo XX, la investigación etnográfica se vuelca hacia la propia sociedad del analista, con enfoques "multiculturalistas" que se sustentan en el pluralismo y la diversidad cultural. En el marco de este proceso que Burgess (1984) denomina "la vuelta a casa de la etnografía", ésta quedará definida por su carácter multimétodo, basado en la utilización de diversos instrumentos de recolección de informaciones (observación participante, entrevista en profundidad, análisis documental), la orientación hacia la especificidad cultural del fenómeno estudiado y el análisis en profundidad de pocos casos (véase apartado 1.5).

Con los trabajos de Park y Burguess la etnografía se reorienta hacia la propia sociedad. Su mirada se encauza hacia la diversidad ciudadana, a través del análisis de todas las otredades urbanas (tribus, ghettos, nacionalidades, etnias, religiones, subgrupos). Herbert Blumer, por su parte, siguiendo la línea de pensamiento de George Mead, desarrolla el interaccionismo simbólico, orientado a comprender toda situación social desde la visión e interpretación del propio actor en interacción. Desde esta perspectiva, los individuos, en tanto sujetos interactuantes y autointeractuantes, deben interpretar el mundo para poder conducirse en él. El interaccionismo simbólico investiga este proceso por el cual los sujetos desarrollan estrategias a partir de las interpretaciones que realizan de su propia experiencia. El método consiste en "asumir el papel del actor y ver su mundo desde su punto de vista" (Blumer 1969: 73). En este proceso investigativo interactivo también se recrea la identidad del investigador.

Pero en la sociología, estas tradiciones que podríamos llamar en cierto sentido protocualitativas perdieron relevancia especialmente luego de la Segunda Guerra Mundial, cuando el maridaje entre la lógica hipotético-deductiva y el *survey*, en el marco del consenso ortodoxo de las ciencias sociales, sellaron el destino de gran parte de la investigación social empírica.[8] A partir de los años sesenta, las críticas a la sociología convencional favorecieron la reactivación de la investigación no estándar. En ese contexto se resignificaron algunas prácticas metodológicas de la antropología clásica y de la sociología de Chicago, dotándolas de un fundamento teórico-epistemológico de corte interpretativista que aparentemente no habían tenido en sus orígenes (véase nota 7). En esta resignificación jugaron un papel central los aportes de la hermenéutica, el constructivismo, la fenomenología, la teoría crítica y otros.

Si bien la tradición interpretativa —como el párrafo precedente sugiere— no constituye un bloque monolítico, todas sus variantes comparten la preocupación por elucidar los procesos de construcción de sentido, aunque la conceptualización de este proceso y las propuestas para su comprensión —como se acaba de decir— no conforman un paradigma único. Entre los aspectos comunes se destaca el interés por diferenciar el objeto de las ciencias sociales, poniendo en cuestión la validez universal del modelo de las ciencias naturales. Asimismo, se evidencia la necesidad de contar con esquemas de investigación y análisis propios —y fundamentalmente diferentes— persiguiendo objetivos cognitivos que, sin perder su carácter científico, no busquen necesariamente la medición y la cuantificación de los fenómenos ni el control empírico de enunciados que den cuenta de sus relaciones.[9] Por otra parte, estas posiciones com-

[8] También se ha destacado la influencia de la nueva situación de EE.UU. a partir del fin de la Segunda Guerra Mundial, como actor hegemónico internacional, en el favorecer los estudios de *survey* por sobre los estudios etnográficos de caso. Aparentemente, los primeros habrían sido más funcionales a las necesidades de conocimiento social para su agenda interna y a la imposición de su modelo político-económico a nivel internacional.

[9] Para un tratamiento detallado de las características de los enfoques no estándar véase el apartado 1.5.

parten la convicción sobre la incapacidad del positivismo para capturar la naturaleza y complejidad de la conducta social, entendida como un fenómeno único, irrepetible e irreplicable, cuyo sentido debe ser comprendido en su especificidad.

En la *New School of Social Research*, de Nueva York, el sociólogo alemán Alfred Schutz, siguiendo la línea de pensamiento fenomenológica de Husserl, sostuvo que los individuos constituyen y reconstituyen su mundo de experiencia. Schutz focalizaba sus estudios en los procesos a través de los cuales los sujetos producen interpretaciones que dan forma a lo real en la vida cotidiana. Ésta era interpretada en el marco del sentido común desde el que se producen y organizan las motivaciones y las acciones. De este modo, la mirada fenomenológica se orientó al razonamiento práctico del sentido común en tanto otorgador de significado. Según Schutz (1962), las ciencias sociales están conformadas por constructos de segundo orden, que le dan sentido a los constructos de primer orden de los actores en la vida cotidiana.

Heredera de la FENOMENOLOGÍA de Schutz es la ETNOMETODOLOGÍA de Harold Garfinkel (1967). La preocupación de este autor, discípulo de Parsons, se orienta a la producción, legitimación, cuestionamiento y reproducción del orden social por la actividad interpretativa de los sujetos. Desde la perspectiva etnometodológica, la conducta —que siempre es imaginada como producto de una norma— es descrita y explicada por su referencia a consensos socialmente compartidos. El análisis se focaliza en el proceso de otorgamiento de sentido a partir de normas sociales, que dan como resultado los criterios de normalidad. En términos instrumentales se prioriza el análisis de conversaciones, enmarcadas en el contexto donde los actores resuelven situaciones sociales. La conversación conlleva y constituye las expectativas que subyacen a la interacción social. Una herramienta etnometodológica son los experimentos basados en estímulos disruptores de expectativas, a fin de analizar las respuestas de los sujetos tendientes a readaptar la situación.

La corriente CONSTRUCTIVISTA, proveniente también de la tradición fenomenológica, se opone igualmente al realismo empírico de la mirada positivista, predicando la imposibilidad del conocimiento objetivo. Esta imposibilidad se fundamenta en la existencia de múltiples realidades, construidas desde diferentes perspectivas. Ante la inexistencia de un mundo real único pierden sentido los criterios rígidos para articular un consenso metodológico y la idea de una ciencia acumulativa organizada en un proceso de desarrollo lineal.

Desde la antropología interpretativa, Clifford Geertz (1973), en oposición al modelo estructuralista de Lévi-Strauss, propone la interpretación como alternativa a las explicaciones causales de la cultura. Para Lévi-Strauss, la conducta de los individuos era explicada a partir de la estructura social y el significado era indagado en las reglas constitutivas del sistema organizador de las acciones. Geertz, por el contrario, considera que la DESCRIPCIÓN DENSA constituye la forma adecuada para dar cuenta del proceso de formación de sentido del sujeto.

Otra vertiente dentro de los métodos cualitativos está representada por la estrategia de la TEORÍA FUNDAMENTADA (*grounded theory*), introducida por Glaser y Strauss (1967) como un método para generar teoría de alcance medio a partir de los datos empíricos. La propuesta, basada en una acción dialógica entre datos y teoría, se opone a la "gran teoría"[10] y al método hipotético-deductivo. La relación entre teoría y datos se invierte: éstos pasan de ser la instancia verificadora/falseadora a ser la fuente de la teoría. El protagonismo de los datos dentro del proceso de investigación se desliza entonces desde la etapa de la puesta a prueba al proceso de construcción de la teoría (véase apartado 17.3). La *grounded theory* promueve una metodología fundada en la inducción, en la que el muestreo estadístico es suplantado por el muestreo teórico. Las unidades de observación se seleccionan según criterios formulados desde la perspectiva del investigador, y el tamaño de la muestra se alcanza por saturación (véase apartado 12. 5), es decir, cuando ninguna observación adicional agrega nueva información relevante.

Vinculada a la ESCUELA CRÍTICA DE FRANKFURT (Adorno, Horkheimer, Marcuse, Habermas, entre otros) nace la denominada TEORÍA CRÍTICA, cuyos mentores, a partir del pensamiento marxista, y en algunos casos también freudiano, desarrollan interpretaciones de nuevas formas de dominación, incorporando la valoración política entre los elementos del proceso investigativo. Cuestionan la objetividad del pensamiento moderno y lo interpretan en términos de sus consecuencias políticas. Denuncian a la ciencia social del consenso ortodoxo, que celebraba la sociedad tal como era y limitaba su función a la mera descripción y explicación de ésta (juicios de hecho), y fomentan en cambio una ciencia transformadora de la realidad social.[11] Las posiciones críticas entienden que las ciencias sociales no están para conservar y reproducir las desigualdades sociales, sino para denunciarlas y modificarlas, e incorporan a los valores y a la ideología como orientadores del análisis de la producción y reproducción de las diversas formas de dominación social.

Actualmente existen muchas escuelas que se insertan en la teoría crítica desde diferentes perspectivas, sin tener un consenso único. En las últimas décadas, por ejemplo, dentro de los nuevos paradigmas de la crítica social, se ha desarrollado la crítica epistemológica y metodológica feminista. Ésta se ha caracterizado, en diferentes momentos de su desarrollo, por un cuestionamiento epistemológico al esencialismo, en la medida que éste apunta a la legitimación de las relaciones de dominación entre los géneros anclada en la biología. Esto dio lugar al surgimiento de nuevas categorías de análisis como el "género", que permitía diferenciar las construcciones sociales y culturales del determinismo biológico, y combatir la naturalización de las desigualdades sociales.

[10] En el sentido que Charles Wright Mills (1959) atribuía a Talcott Parsons, especialmente por su teoría general de la acción y del sistema social.

[11] Se trata de un momento histórico en el que se realiza una importante revisión crítica de la idea de ciencia social ortodoxa —pensada y justificada a imagen y semejanza de la física— que por un lado se abrazaba a una supuesta objetividad y neutralidad, mientras que por el otro contribuía a difundir una visión conservadora de la sociedad. Y esta revisión crítica general y generalizada incluyó obviamente a la problemática de los métodos.

Otra importante crítica feminista se orientaba al carácter de verdad necesaria y universal con que el determinismo biológico impregnaba la posición subordinada de las mujeres. A esto se le opuso el carácter contingente e históricamente situado de las relaciones sociales.

Un rasgo de la teoría crítica que se expresa en el feminismo es su traducción en la práctica política, manifestada en la lucha contra el sexismo, contra los valores y las instituciones del patriarcado y contra las relaciones de poder-subordinación entre los géneros.

A principios de los ochenta muchas representantes de esta corriente abandonaron la idea de una gran teoría feminista capaz de generar categorías para explicar al sexismo sin tomar en cuenta la diversidad cultural. En parte, esto fue una respuesta a fuertes críticas internas hacia la teoría hegemónica que fue caracterizada como "de y para mujeres blancas heterosexuales de clase media". Estos debates dieron lugar a la incorporación de perspectivas epistemológicas que favorecieron una comprensión clasista, sexual y étnica de las relaciones sociales, histórica y culturalmente situadas.

Actualmente la mirada se vuelca hacia investigaciones de tipo localista; la construcción académica feminista se orienta cada vez más a las diferencias y las particularidades históricas y culturales, abandonando la idea de una teoría, una epistemología o un método feministas. Como sostienen Fraser y Nicholson (1990), tal vez sería mejor hablar de la práctica política feminista contemporánea en plural, como la práctica de los feminismos.

La reactivación del interés por la investigación cualitativa, a partir de la década de 1960, también implicó que gran parte del debate metodológico se trasladara desde la oposición cuantitativo/cualitativo hacia el interior de los métodos cualitativos. Schwandt (1994: 130) identifica cuatro ejes en el debate intracualitativo:

a) los criterios y la objetividad, centrado en cuál es el fundamento de las interpretaciones;
b) el alcance crítico o compromiso político, que lleva a una distinción entre teoría descriptiva y prescriptiva;
c) la autoridad del investigador a partir de la legitimidad de la interpretación, y
d) la confusión de las demandas psicológicas y epistemológicas.

Una de las principales críticas a los métodos no estándar se ha dirigido a su supuesta falta de rigor y precisión, así como su incapacidad para generalizar los resultados. Como respuesta a estas críticas, y a partir del cuestionamiento de los criterios positivistas de objetividad, validez, fiabilidad y generalización, se instaló en la agenda cualitativa la búsqueda de instrumentos conceptuales y criterios alternativos que fuesen más adecuados para juzgar sus procedimientos y sus productos. Guba y Lincoln (1985), por ejemplo, han propuesto las nociones de credibilidad y transferibilidad; Adler y Adler (1994), las de autenticidad y verosimilitud, y con respecto a la posibilidad de generalización, Patton (1986) y Sykes (1991) han defendido la idea de extrapolación razonable (véase apartado 10.5). Sadowski y Barroso (2002), por su parte, señalan que a pesar del gran esfuerzo realizado en los últimos veinte años con respecto a esta cuestión, aún no se ha llegado a establecer un consenso acerca de los criterios de calidad en los estudios cualitativos; a su juicio, ni siquiera se ha alcanzado un acuerdo sobre la necesidad/pertinencia de que exista un consenso sobre tales criterios.

Tal vez esto se deba, al menos en parte, al hecho de que desde la década de 1990 —en la que Denzin y Lincoln (1994) identifican un "quinto momento" en el desarrollo de los métodos cualitativos— coexisten una multiplicidad de concepciones que tienen en común la perspectiva del sujeto: neoestructuralismo, interaccionismo simbólico, antropología cultural y cognitiva, feminismo, etnometodología, fenomenología, teoría crítica, estudios culturales, etcétera. Para Denzin (1994), esta variedad de enfoques cualitativos responde a la realidad de un mundo posmoderno, múltiple y fragmentado, en el que no es posible ninguna generalización y donde coexisten diferentes concepciones competitivas, sin predominio de unas sobre otras (véase apartado 1.5). Esta multiplicidad de perspectivas, por otra parte, se da en un contexto en el que los límites entre los estilos cuantitativo y cualitativo se han vuelto imprecisos —si es que alguna vez fueron muy precisos— y se generan incertidumbres en cada uno de ellos (Hamilton 1994). Para Valles (1997) la característica del presente, en términos de la metodología de las ciencias sociales, está dada por el repliegue, por la autocrítica de cada uno de los dos grandes modelos de investigación. En línea con esto, desde hace ya muchos años han empezado a difundirse los intentos de integración de ambos tipos de investigación, e incluso los de superación del debate *tout court.*

Una figura destacada y pionera en los intentos de superación del debate cualitativo/cuantitativo fue la de Pierre Bourdieu, heredero de una tradición sociológica clásica cuyas variadas influencias, en cierto sentido, buscó sintetizar de un modo original. Una particularidad de Bourdieu fue su carácter polifacético: él se ha movido con relativa comodidad por los senderos de la investigación empírica, de la teorización sociológica, y de la reflexión epistemológica y metodológica —ámbitos que la profesionalización y la especialización de las ciencias sociales han tendido a mantener relativamente aislados entre sí. Tal vez haya sido esta particularidad lo que le permitió promover la superación de las típicas antinomias epistemológicas y metodológicas de las ciencias sociales: explicación/comprensión, subjetivo/objetivo, micro/macro, estándar/no estándar, etcétera. Desde el punto de vista operativo, Bourdieu buscó integrar prácticas habitualmente reconocidas como cualitativas (por ejemplo, la observación participante), con otras cuantitativas (encuesta por muestreo y análisis estadístico).

Pero más allá de estos intentos de superación del debate cualitativo/cuantitativo, entre los cuales el de Bourdieu representa tan sólo un ejemplo, en los últimos años se han propuesto diversas formas de articulación o integración entre las diferentes perspectivas metodológicas. Éstas, que genéricamente se conocen como triangulación metodológica, serán el objeto de la siguiente sección.

2.4. Las propuestas de superación del debate cualitativo-cuantitativo: la triangulación metodológica y sus límites

Como se acaba de indicar, cada vez se han vuelto más recurrentes las críticas a la distinción cualitativo/cuantitativo. A partir de una revisión de la literatura sobre el tema, Chiesi (2002) concluye que en la actualidad la mayoría de los autores la considera banal o simplemente incorrecta. Cardano (1991), por su parte, indica que es ambigua, y que en todas las tradiciones teóricas de las ciencias sociales se han desarrollado técnicas que podrían considerarse cualitativas o cuantitativas. Para Statera (1992) se trata de una cuestión meramente retórica, y Campelli (1991) opina que todos los actos de investigación empírica implican una combinación de cualidad y cantidad.

En línea con estas miradas críticas ha surgido una serie de propuestas tendientes a integrar ambos tipos de estrategias. Ellas ya forman parte del patrimonio de conocimiento metodológico de las ciencias sociales y aparecen reflejadas cada vez más en los manuales especializados. Como plantea Ruiz Olabuénaga (1996), en estos textos pareciera sostenerse una posición de relativo consenso, ligada al argumento técnico que hemos presentado en la primera sección de este capítulo, en torno de dos cuestiones: 1) ambos métodos son igualmente válidos, aunque por sus características resultan recomendables en casos distintos, y 2) ambos métodos no son incompatibles, hecho del que deriva la posibilidad de integración siempre que esto repercuta en un mejor abordaje del problema en cuestión. Este último postulado, justamente, remite a las ideas actuales acerca de la triangulación.[12]

Para Bryman (2004), una de las primeras referencias a la triangulación se encuentra en la idea de *unobtrusive methods* propuesta por Webb *et al.* (1966), quienes además habrían sido los primeros en usar la palabra 'triangulación' en las ciencias sociales. Sin embargo, la idea —al menos en un sentido metafórico— ya había sido desarrollada por Campbell y Fiske (1959) (véase también el apartado 7.5), entre otros, para dar cuenta de la convergencia de diferentes mediciones en la determinación de un mismo constructo, evitando de este modo las limitaciones de una única operativización. Sólo que en estas obras la práctica había sido designada recurriendo a otras expresiones, como por ejemplo la de validez convergente. Sin embargo, la noción de Campbell y Fiske era fiel al significado que la triangulación tenía en las disciplinas de las cuales fue tomado el término (la navegación y la agrimensura): esencialmente, éste hace referencia a la determinación de la posición de un punto a partir de la intersección de dos rectas trazadas desde otros dos puntos cuyas posiciones son conocidas.

Denzin (1970) fue uno de los principales responsables de la difusión de la idea de triangulación, distinguiendo además cuatro formas: de datos, de investigadores, teórica y metodológica. A su vez, promovió una diferenciación entre la triangulación intramétodo (*within-method*) e intermétodo (*between-method*).

[12] Para un tratamiento más exhaustivo del tema de la triangulación en las ciencias sociales, sus orígenes, desarrollo histórico, modalidades, críticas, etcétera, véase Piovani *et al.* (2006).

Knafl y Breitmayer (1989) señalan que los principales objetivos de la triangulación en las ciencias sociales son la convergencia y la completitud: como medio de validación convergente o como modo de alcanzar una comprensión más acabada de un mismo fenómeno, desde distintos ángulos.

A pesar de sus atractivos, la idea de triangulación también ha sido objetada, y desde diversos puntos de vista. Bryman (2004) enumera algunas de las críticas más frecuentes: sus presupuestos ligados a cierto realismo ingenuo y la asunción de que datos provenientes de distintos métodos puedan ser comparados inequívocamente (e incluso considerados equivalentes). Massey (1999) realiza una crítica más radical; para él, los fundamentos conceptuales de la triangulación son esencialmente inconsistentes.

De todos los tipos de triangulación, la metodológica es probablemente la que ha adquirido mayor difusión y popularidad, al punto que se han acuñado nuevas expresiones —como "investigación multimétodo"— para dar cuenta de ella. Según Bryman (2004), ésta implica el uso conjunto de dos o más métodos para abordar un mismo problema o problemas diferentes pero estrechamente relacionados. El principal argumento a su favor (que Massey considera erróneo desde un punto de vista lógico) es que de este modo se aumenta la confianza en los resultados de una investigación.

La propuesta de triangulación metodológica ha encontrado oposición especialmente entre aquellos más apegados a la explicación epistemológica de las diferencias entre métodos cualitativos y cuantitativos. No obstante, como indica Bryman (2004), la mayoría de los investigadores ha adoptado posiciones más pragmáticas: incluso admitiendo que las distintas formas de investigación conllevan compromisos epistemológicos y ontológicos contrapuestos, ha aceptado la idea de que la combinación de sus respectivas fortalezas puede generar beneficios para la investigación. Pero esta última afirmación puede resultar en cierto sentido voluntarista: la combinación de métodos no siempre tiene un efecto compensatorio de sus desventajas y potenciador de sus ventajas; en ocasiones se refuerzan las limitaciones y se multiplican los sesgos. En este sentido, Massey (1999) afirma que en la actualidad resulta perentorio identificar los modos adecuados e inadecuados de combinar los métodos, y Jick (1979), por su parte, señala la necesidad de desarrollar pautas concretas acerca de cómo combinarlos.

En todo caso, las formas de articulación prevalecientes hasta el presente se han mantenido relativamente apegadas a estrategias convencionales: empleo de la perspectiva cualitativa en fases exploratorias, para la posterior realización de estudios cuantitativos de contrastación de hipótesis; utilización de la investigación cuantitativa para establecer regularidades y tipos sociales que luego se exploran en detalle a través de estrategias cualitativas; uso conjunto de ambos métodos para indagar las relaciones entre casos "micro" y procesos "macro" (Bryman 1988).

Ante estas consideraciones, y a pesar de la popularidad que la triangulación ha adquirido, queda claro que ésta aún no ha logrado cerrar la brecha entre métodos estándar y no estándar, ni siquiera en términos de su posible articulación/integración.

Por otra parte, se corre el riesgo de que la triangulación se convierta en una nueva moda metodológica: como advierte Bryman (2004), actualmente pareciera haber una tendencia —para él injustificada— a creer acríticamente que una investigación multimétodo es superior a una monométodo. Así como durante mucho tiempo —y en algunos contextos aún hoy— se recurría a las técnicas estadísticas como fetiches, sin controlar concienzudamente su utilización, hay algunos elementos que permiten pensar que algo análogo ha empezado a suceder con la triangulación: en muchos casos se recurre a ella de modo ritualista, olvidando que se trata simplemente de una metáfora (como indica Massey 1999), y descuidado la reflexión crítica acerca de los problemas metodológicos.

CAPÍTULO 3

MÉTODO, METODOLOGÍA, TÉCNICAS

3.1. El origen griego del término 'método'

Como a muchos otros términos, al término 'MÉTODO' se le atribuyen significados diferentes no sólo en el lenguaje científico sino también en el lenguaje ordinario. Una breve reseña de los diccionarios monolingües de varios idiomas permite afirmar que la gama de acepciones en el lenguaje ordinario corresponde aproximadamente a la del lenguaje científico.[1]

La acepción que se encuentra como primera en casi todos los diccionarios es la más cercana al significado originario griego (véase más adelante) y también la que se propone aquí como la preferible en el lenguaje científico: método como camino para conseguir un fin. Se encuentran también regularmente acepciones cercanas a las que se consideran "bajas" en el lenguaje científico: método como procedimiento, técnica. Otras acepciones que se presentan en uno u otro diccionario monolingüe son: criterio, forma, costumbre, uso corriente, praxis, y también artificio, engaño.

El prefijo met- (meta) denuncia el origen griego del término. Ese prefijo aparece en muchos términos académicos de las lenguas occidentales en sus diferentes significados: como "más allá" en *metafísica, metástasis, metalenguaje*; como "en lugar de" en *metáfora, metatesis, metonimia, metamorfosis, metempsicosis.* En combinación con el sustantivo οδος (camino), el prefijo met- asumía otro de sus significados principales ("con"). El compuesto μεθοδος significaba por lo tanto "camino con [el cual]". El significado en el lenguaje ordinario griego clásico (sucesión de actos tendientes a conseguir un fin) quedó fiel a la etimología del término.

Después de casi veinte siglos, en la definición que da del término la *Logique de Port Royal* (*"ars bene disponendi seriem plurimarum cogitationum"*:

[1] Parece, en efecto, que la tesis —de Bachelard (1934, 1938) y de muchos otros— de una separación neta (la *coupure épistemologique*) entre lenguaje ordinario y lenguaje científico sea falsa como descripción de la situación —donde se encuentra a menudo una plena continuidad— y no oportuna como prescripción. Se prefiere la tesis de Schutz (1953) según la cual el científico social debe tratar de sentar sus términos en el lenguaje ordinario, o al menos de hacerlos comprensibles a sus objetos, que —a diferencia de las ciencias físicas— también son sujetos.

Arnauld y Nicole 1662) la idea de una sucesión de pasos sobrevive y se conjuga con la del arte (entendida en el sentido de habilidad).

3.2. La visión clásica del método

Primero en la tradición filosófica occidental, Descartes proclama el papel central del método en la actividad intelectual (*Discours de la Méthode*, 1637). Él mismo ya había definido el método como "reglas ciertas y fáciles que cualquiera fuera el que las observara con exactitud le sería imposible tomar lo falso por verdadero, sin requerir inútilmente esfuerzos de la mente, pero aumentando siempre gradualmente el [propio] saber, y lo conducirían al conocimiento verdadero de todo aquello que fuese capaz de conocer" (Cuarta *Regula ad directionem ingenii*, 1628).

Esta naturaleza universal del método había sido subrayada algunos años antes por Francis Bacon en el aforismo 122 del *Novum Organum* (1620): "Nuestro método de investigación pone casi a la par todos los intelectos, porque deja poco espacio a las capacidades individuales, pero las liga con reglas muy sólidas y con demostraciones".

Relacionando pacientemente las declaraciones de dos pensadores considerados en la raíz de dos direcciones contrapuestas en el pensamiento filosófico-científico contemporáneo (racionalismo y empirismo), impacta la extraordinaria similitud de sus conceptos de método:

- no estamos en presencia de una cosa esotérica, especializada. Es más, las reglas a seguir son fáciles, automáticas ("sin requerir esfuerzos de la mente"), al alcance de todos ("de cualquiera", "pone a la par todos los intelectos");
- conocimientos especializados, capacidades, iniciativas personales, no sólo no son requeridos, sino que son rechazados ("deja poco espacio a las capacidades individuales, pero las liga a reglas muy sólidas y con demostraciones"). Las reglas se imponen para todos ("reglas ciertas... observadas con exactitud").

Si se siguen fielmente —agrega Descartes—, las reglas de este método "conducirán al conocimiento verdadero".[2]

En los dos o tres siglos siguientes, por efecto de los éxitos que la matematización de la ciencia va consiguiendo con Kepler, Galileo, Newton, Boyle, Lavoisier, esta idea de "un método matemático universal, aplicable en todos los campos del conocimiento científico" asume connotaciones "pitagórico-platónicas" (véase Berka 1963: 8) que no tenía en Bacon y tenía sólo potencialmente

[2] Carlos Strasser ha resaltado (1979: 19) otros rasgos que Bacon y Descartes tienen en común: el desprecio por todos los conocimientos que la sabiduría humana había acumulado antes de ellos, y el énfasis sobre la utilidad práctica de los descubrimientos.

en Descartes. Se viene explicitando cada vez más la idea —implícita en el concepto griego de επιστημη— de que verdad y certitud son conjugables, de que se puede, por lo tanto, hablar de la realidad que nos circunda sin renunciar a la certeza de lograr el nivel de veracidad que sólo los matemáticos pueden garantizar: una ambición que será la bandera del empirismo lógico —como muestra la autodenominación del movimiento.

La idea de un programa que establece con anticipación una serie no modificable de operaciones que, de ser practicadas, garantizan el conocimiento científico sobre cualquier argumento, es todavía hoy uno de los significados que prevalece en la expresión 'método científico'. Pero —visto también las dificultades encontradas por los empiristas lógicos para transferir aquellas ideas desde el plano de los enunciados de principio a sus aplicaciones— tal idea queda rezagada en lugar de ocupar triunfalmente el escenario como en los tiempos de Bacon y Descartes. Lo confirma entre otras cosas la creciente escasez de textos que propongan especificar la mencionada serie no modificable de operaciones.

La más reciente —a nuestro entender— de dichas listas de operaciones que tiene, sin entrar en detalles, una cierta pretensión de completud ha sido publicada por Theodorson y Theodorson en la voz *scientific method* en su *Diccionario de sociología* (1970: 370). Dice así:

> La aplicación del método científico a un problema comporta los pasos siguientes:
>
> - primero: se define el problema;
> - segundo: el problema se formula en los términos de un particular cuadro teórico;
> - tercero: se imaginan una o más hipótesis relativas al problema, utilizando los principios teóricos ya aceptados;
> - cuarto: se determina el procedimiento a utilizar en la recolección de datos para controlar la hipótesis;
> - quinto: se recolectan los datos;
> - sexto: se analizan los datos para averiguar si la hipótesis es verificada o rechazada.
>
> En fin, las conclusiones del estudio son relacionadas con el corpus precedente de la teoría que se modifica para ajustarlo a los nuevos resultados.

Como el lector habrá observado, se trata de una descripción puntual del así llamado "método hipotético-deductivo" elaborado a partir de la reflexión filosófica acerca de la física del siglo XVII en adelante (Losee 1972: caps. 7-9). La convicción de que éste sea el único método posible para la ciencia ha sido muy difundida incluso en las ciencias sociales.[3] Hasta los años cincuenta no era para nada raro encontrarse con declaraciones como la siguiente, del antropólogo Nadel: "Hay un único método científico, a pesar de ser practicado con diversos

[3] Lo relevan, entre otros, Strasser (1979), Pérez Ransanz (1995).

niveles de rigor y coherencia, y a este respecto la física y la química han alcanzado los mejores resultados. Cada investigación está vinculada a este método, y ningún otro es concebible" (1949/1972: 227).[4] En epistemología, la tesis de la unidad del método científico ha sido afirmada también recientemente, en una forma más cauta: "Se puede razonablemente afirmar que las diversas disciplinas científicas tienen el mismo método en lo que se refiere a los procedimientos o al complejo de reglas que las integran" (Pera 1978: 11).

3.3. Críticas a la visión clásica

Pero no todos están de acuerdo. Dalton ha observado polémicamente: "Se presume que una secuencia invariable: formulación de hipótesis / control / confirmación sea compartida por todos los científicos, que todos concuerden en que ésta es la única vía al conocimiento; en resumen, que existe un único método científico" (1964: 59). En cambio "nos podemos preguntar —observa Howard Becker— si los metodólogos, los guardianes institucionales de la metodología, afrontan realmente el entero espectro de las cuestiones metodológicas relevantes para la sociología, o si se limitan a un subconjunto no aleatorio (como dirían ellos) de tales cuestiones" (1970: 3). En el mismo espíritu, el sociólogo y epistemólogo polaco Mokrzycki sostiene: "Los fundamentos de este 'método' se encuentran fuera de la sociología, carentes de contacto con el pensamiento sociológico. La 'metodología de las ciencias sociales' se ha vuelto una cinta de transmisión que distribuye a los sociólogos el 'método científico', es decir, las ideas de aquellos autores que pasan por ser expertos en el tema" (1983: 72). "Muchos de nosotros —agrega Dalton— aceptamos el 'método científico' porque estamos convencidos de que se ha desarrollado en las ciencias naturales... Pero en las ciencias naturales no se profesa deferencia a este modelo como lo hacemos nosotros" (1964: 59).

La vivacidad de estas reacciones se explica por su relación con un período, no lejano, en el cual la orientación cientificista dominaba en las ciencias sociales: se sostenía —como se ha visto en los pasajes citados— no sólo que la ciencia tiene un solo método, sino que se trata de aquel método santificado por los éxitos de las ciencias físicas. En los años recientes, esta idea de la unidad del método ha sido criticada. "La idea misma de 'un método' acrecienta la confianza en los propios resultados y reduce la predisposición a preguntarse si los procedimientos consolidados tienen sentido en el caso particular" (Kriz 1988: 184). Para el epistemólogo francés Haroun Jamous, en la raíz de esta orientación se encuentra la "repulsión por la incertidumbre, [...] la perenne necesidad de riendas estables y definitivas que puedan dispensarnos de recurrir a aquel

[4] Declaraciones tan categóricas se pueden encontrar en las obras del psicólogo conductista Skinner (1953: 5), del sociólogo operacionista Lundberg (1938: 191-192), del antropólogo Murdock (1949: 183) y —naturalmente— también en las obras de los epistemólogos neopositivistas (Neurath 1931-32: 407; Hempel 1935/1948: 382) o cercanos al neopositivismo (Kemeny 1959/1972: 27; Rudner 1966/1968: 18-9).

aporte individual e incierto que, justamente por ser difícilmente formalizable, es tal vez indispensable en toda obra creativa" (1968: 27).

Vinculada a la precedente, una segunda línea crítica señala el peligro de que una particular serie de procedimientos, identificada con "el método científico", se reifique y se vuelva un fin en sí.[5] "Si el criterio es el respeto a un sistema de reglas —observa maliciosamente Lecuyer— incluso una guía telefónica es una buena investigación científica" (1968: 124). Y el sociólogo alemán Weigert nota irónicamente que "la característica de la magia es la atribución de eficacia a rituales minuciosamente observados sin indagar el nexo causal entro rito y efecto" (1970: 116). En el mismo espíritu irónico Merton observaba que la secuencia rigurosa de actos aparece invariablemente en el momento de presentar los informes de investigación (1949: 506) —lo que recuerda las punzantes observaciones de Fleck sobre el proceso de racionalización *ex post* que sufren los procedimientos experimentales (1935: § 4.2).

Aun compartiendo muchas de estas críticas a la tesis de la naturaleza única y fija del método científico, no se puede suscribir a la *boutade* de los estadísticos Wallis y Roberts, para quienes "no hay nada que pueda llamarse método científico. Es decir, no existen procedimientos que indiquen al científico cómo empezar, cómo proseguir, qué conclusiones alcanzar" (1956: 5). Aunque raramente explicitada en todo detalle, en la consciencia colectiva de los investigadores en ciencias sociales una cierta idea de aquel método científico existe: se la encuentra por ejemplo en el recurso de muchos autores a palabras clave como 'hipótesis', 'verificación', 'experimental', que cumplen una función de legitimación de su propia obra.

Pero "la calidad de un científico —recuerda Toulmin— se demuestra menos por su fidelidad a un 'método' universal que por su sensibilidad a las exigencias específicas de un problema" (1972, vol. I: 150). En sus comentarios a una compilación de ensayos clásicos de investigación social, Madge sostiene: "Si es justo pedir a un investigador que indique lo más precisamente posible aquello que pretende hacer antes de iniciar el trabajo, no se sigue necesariamente que quienes hacen exactamente lo que habían establecido sean los mejores investigadores [...] Las virtudes del buen investigador son la flexibilidad, la prontitud para comprender una nueva situación y la capacidad de aprovechar las ocasiones imprevistas" (1962/1966: 290). En sí, la fidelidad a los procedimientos codificados no ofrece ninguna garantía: "dado un procedimiento adecuado, es posible reconocer investigaciones consideradas pseudocientíficas que lo satisfacen" (se trata de la "paradoja del procedimiento" resaltada por Pera 1991: 27). "Sea un buen artesano, que elige vez por vez qué procedimiento seguir", recomienda Mills, y agrega una invitación: "pedimos a quien haya producido trabajos de calidad contarnos en detalle cómo ha operado: sólo en estas conversaciones con investigadores expertos los jóvenes podrán adquirir una útil sensibilidad metodológica" (1959: 224 y 228).

Son muchos los que dicen que el método es sobre todo una elección. "La cuestión metodológica propiamente dicha es la elección de la técnica en fun-

[5] Véanse, por ejemplo, Kaplan (1964), Parisi y Castelfranchi (1978).

ción de la naturaleza del tratamiento que cada técnica impone a su objeto" (Bourdieu *et al.* 1968: 59). "En la toma de decisiones consiste en definitiva la investigación científica" (Samaja 2001: 16). "Cada investigación es un largo sendero con muchas bifurcaciones en las que se debe tomar una decisión [...] Ninguna regla, ningún algoritmo puede decir cuál es la decisión justa [...] Cuanto más el investigador conciba al método como una secuencia rígida de pasos, más decisiones tomará sin reflexionar y sin darse cuenta" (Kriz 1988: 81 y 131). Además, tener que escoger no significa necesariamente el deber afrontar cada problema con una sola técnica: "Si se elige una sola técnica, los resultados no pueden ser considerados independientes de la técnica elegida, y por lo tanto no son estables: pueden de hecho ser contradichos por resultados obtenidos con otras técnicas" (Van Meter 1994: 23).

Si la investigación debe ser "un proceso cognitivo más que un simple proceso de validación de ideas ya formuladas" (Bailyn 1977: 101), es necesario "saber mantener la tensión entre la necesaria función de reducción de la complejidad fenoménica y la apertura a dimensiones que permitan aumentar la complejidad de los esquemas conceptuales" (Crespi 1985: 343).

El método es por lo tanto algo mucho más complejo que una simple secuencia unidimensional de pasos. "Sin dudas —observa Polanyi— el científico procede en modo metódico. Pero su método es como las máximas de un arte que aplica de manera original a sus problemas" (1958: 311). Su obra es "muy similar a una creación artística... pero no es un arte como la escultura y la pintura, en las cuales uno es libre de obtener cualquier cosa del material crudo. Es un arte como la arquitectura, en la cual se puede mostrar creatividad trabajando con materiales crudos caracterizados por limitadas propiedades ingenieriles, y para comitentes con presupuestos vinculados a objetivos precisos" (Davis 1964: 267-8). La gestión de recursos en presencia de vínculos es un aspecto característico del concepto de método; luego volveremos a esto.

Por el momento se puede concluir que la visión cartesiano-baconiana del método, que hasta hace algunos decenios no se ponía en tela de juicio, hoy en día genera perplejidad. No parece haber escandalizado a nadie Marcello Pera cuando ha descrito la "paradoja del método": "la ciencia se caracteriza por su método, pero una caracterización precisa del método destruye la ciencia" (1991: 39).

3.4. Método y metodología

Sin embargo, este cambio de perspectiva no ha dejado todavía ninguna huella en el lenguaje. Una consecuencia inevitable del hecho de que la orientación prevaleciente a lo largo del tiempo en la filosofía de las ciencias sociales reificase el método en una sucesión de pasos y procedimientos ha sido que el término 'método' no sólo designe un modo particular de cumplir aquellos pasos ("método experimental", "método hipotético-deductivo"), sino también algún paso específico, e incluso los instrumentos operativos que permiten cumplir los varios pasos (es decir, las técnicas: véase apartado 3.5). En esta

acepción el término ha desarrollado un plural (para designar el complejo de técnicas, el grupo de técnicas de un cierto tipo) que no se justificaba en ninguna de las acepciones examinadas en las secciones precedentes. El uso de 'método/métodos' en el sentido de 'técnica/técnicas' ha sido y es endémico en las disciplinas más orientadas a la investigación empírica, que están más frecuentemente en contacto con instrumentos operativos.

Kaplan distingue los métodos de las técnicas por cuanto los primeros valen de modo bastante general en todas las ciencias o en una parte importante de ellas (1964: 23). Se encuentra, en efecto, en el uso una cierta estratificación jerárquica en el sentido indicado por Kaplan (los métodos son más generales y más nobles que las técnicas), pero ningún otro autor —por lo que sabemos— la ha expresado.

Se ha hablado de la indebida expansión semántica del término 'método'. Pero un abuso todavía más grave, como veremos en breve, lo sufre el término 'METODOLOGÍA'. En docenas de términos científicos —muchos de los cuales se han pasado al lenguaje ordinario— el sufijo 'logía' está utilizado en el sentido de "discurso sobre", "estudio de". Por lo tanto, 'metodología' debería ser el discurso, estudio, reflexión sobre el método. Empero en la investigación social norteamericana —cuyos cultores generalmente no tienen alguna idea de las raíces griegas del lenguaje científico— se ha difundido desde hace tiempo (y ha sido implacablemente exportada a otros países) la costumbre de llamar *methodology* a la técnica. Es éste el abuso terminológico que se señalaba más arriba. Se comprenden las protestas del interaccionista Blumer hacia esta "asombrosa inclinación a identificar la metodología con una limitada porción de su objeto de estudio" (1969: 22; véase también Sartori 1984: 9).

Sin embargo, hay autores que al definir el término muestran que no han olvidado las raíces griegas de la terminología científica,[6] pero son una minoría. La definición más cercana al espíritu de este texto se debe a Lazarsfeld: "la metodología examina las investigaciones para explicitar los procedimientos que fueron usados, los supuestos subyacentes, y los modos explicativos ofrecidos" (Lazarsfeld *et al.* 1972: xi). Lazarsfeld sintetiza su punto de vista con esta analogía: "la poesía es emoción a la cual se vuelve con ánimo tranquilo. Considero la metodología un volver al trabajo creativo con el mismo estado de ánimo" (1959/1967: 186). Aun no habiendo nunca definido directamente el término 'método', Lazarsfeld muestra en estos pasajes concebirlo como una actividad creativa, a la par de Mills, Polanyi, Davis y otros.[7]

A propósito de la metodología en su sentido propio, se discute si ella debe ser una disciplina prescriptiva o descriptiva. Por la segunda alternativa se ha pronunciado con fuerza Dewey (1938); otros han resaltado que el estudio em-

[6] Por ejemplo Novak: "la metodología se encarga del análisis de los procedimientos científicos y de los instrumentos de investigación" (1976: xv); Smelser: "metodología es la evaluación crítica de las actividades de investigación en relación con los estándares científicos" (1976: 3); Stoetzel: "la metodología puede definirse como el conocimiento que resulta de una reflexión sobre los aspectos empíricos de la investigación" (1965/1969: 10).

[7] Véanse los pasajes de estos autores citados arriba.

pírico de cómo trabajan los científicos es una tarea de la historia y de la sociología de la ciencia; la metodología debe ser prescriptiva. "La metodología —escribe el epistemólogo italiano Bruschi— implica una concepción lógico-racional de la ciencia [...] El aspecto normativo, que en la filosofía de la ciencia puede ser latente, aquí es manifiesto y directo [...] El metodólogo declara aquello que el investigador debe hacer para obtener ciencia" (1991: 38-9).

Se puede aceptar la idea de que la metodología sea, en última instancia, también una disciplina normativa, en cuanto da indicaciones acerca de lo que es oportuno hacer en cada situación específica. Pero tales indicaciones no deben basarse en un modelo abstracto y general de ciencia, sino en aquello que la historia y la sociología de la ciencia nos refieren. Por ejemplo, a la luz del hecho de que una cierta técnica, perfectamente adecuada en abstracto, sea en realidad muchas veces portadora de graves sesgos, v. g. a causa de una insuficiente consideración de los procesos psico-socio-antropológicos que su aplicación tiende a poner en movimiento entre los sujetos observados y/o entre los observadores.

La contraposición descriptivo/prescriptivo encuentra un paralelo en la distinción entre los dos significados que el término 'metodología' tiene en el lenguaje corriente de la investigación social —significados que corresponden a los dos roles del metodólogo profesional. Cuando estudia y enseña, el metodólogo debe tener una orientación descriptiva, es decir, debe estar abierto a aprender de las experiencias de otros investigadores evaluándolas sin preconceptos, y dispuesto a referir en modo sistemático y sintético aquello que ha aprendido. Cuando utiliza sus propias competencias, capacidad y experiencia al servicio de una investigación, suya o de otros, el metodólogo no puede ser otra cosa que prescriptivo, en cuanto debe elegir los instrumentos a utilizar y cómo utilizarlos. En este sentido, será mejor que en sus prescripciones tenga en cuenta todo aquello que ha aprendido en su otro rol.

3.5. Las técnicas

También el término 'TÉCNICA' deriva del griego, en el que τεχνη designa una capacidad artística, no aquella individual e irrepetible del genio, sino aquella más doméstica del artesano, transmisible de padre a hijo (en efecto, el término tiene un probable parentesco con τεκνον, hijo).

La bella definición del sociólogo italiano Gallino ("complejo más o menos codificado de normas y formas de proceder reconocido por una colectividad, transmitido o transmisible por aprendizaje, elaborado con el fin de desarrollar una actividad manual e intelectual de carácter recurrente [...] Un procedimiento extemporáneo, que no se reitera ni sufre alguna forma de codificación, no es una técnica aun si resulta por una vez excepcionalmente ingenioso y eficaz": 1978, 712-3) muestra que muchos elementos del significado original griego se han mantenido al menos en uno de los significados centrales del término.

Gallino toma, en efecto, el significado antropológico, muy difundido también en el lenguaje ordinario. Pero junto a éste se pueden distinguir al menos otros dos, que tienen en común una relación instrumental entre su referente y la ciencia, pero se distinguen por la dirección de tal relación. En un segundo significado, una técnica se sirve del conocimiento adquirido por las ciencias sobre la realidad para modificar algunos aspectos de ésta. En un tercer significado, es una ciencia la que se sirve de técnicas para conocer mejor algún aspecto de su objeto. En ambos conceptos, y en particular en el último —él que aquí nos interesa—, se encuentran presentes aspectos de la intensión del concepto ilustrado por Gallino.

En el ámbito de la medicina, la clínica se sirve del conocimiento adquirido por la fisiología y la patología —que son ciencias— para curar a los enfermos, y es por lo tanto una técnica en el segundo significado. Una situación análoga se da en la ingeniería civil, cuando se sirve de los conocimientos de estadística y dinámica para construir edificios, rutas, puentes. Es cierto —como hace notar el físico italiano Cini— que "la tecnología proporciona a la investigación científica contribuciones originales" (1990: xiv); pero la distinción es analítica: si en el ámbito de la investigación tecnológica se adquieren nuevos conocimientos sobre la realidad que se desea modificar, en este momento se está haciendo ciencia. Además, decir que la técnica (en el segundo significado) aplica los conocimientos científicos para intervenir sobre la realidad no significa disminuirla; somos todos conscientes del hecho de que "la invención de un instrumento muchas veces ha marcado una época" (Jevons 1874: 272).

Las técnicas de las que hablamos aquí son instrumentales a los fines cognoscitivos de las diversas ciencias. Son "los procedimientos específicos empleados en una ciencia, o por un particular género de investigaciones dentro de una ciencia [...] Son las formas de cumplimentar el trabajo de una ciencia que son considerados, por razones más o menos convincentes, aceptables por una comunidad científica. El adiestramiento científico tiene que ver en gran medida con la adquisición del dominio de las técnicas" (Kaplan 1964: 19); en gran medida pero no enteramente, como advierte Lazarsfeld (1959/1967: 229) —y muchos, incluido Kaplan, estarían de acuerdo.

La relación entre método y técnica (en el significado analizado en esta sección) puede ser clarificada con una imagen. Cada investigación tiene un objetivo cognoscitivo: es decir, busca mejorar, profundizar, articular el conocimiento sobre un cierto argumento. Podemos imaginar este objetivo como un claro en una selva: se confía que puede ser alcanzado, se espera alcanzarlo, pero no se sabe exactamente dónde está, y menos aún se tienen claras las ideas sobre cómo llegar. Podemos también imaginar que el investigador (o quien toma las decisiones sobre cómo proceder) parte de un punto cualquiera en el confín de la selva. Raramente está totalmente aislado, y raramente es plenamente libre; generalmente tiene ayudas y también limitaciones. Dispone de fondos públicos o privados, y casi siempre tiene colaboradores más o menos expertos en varias tareas. Puede recurrir, si le parece oportuno, a entes especializados en la recolección de información y/o en el análisis de datos. Entre sus recursos se encuentra también el bagaje de conocimientos derivados de experiencias precedentes de investigación, propias y ajenas. La limitación principal está dada por el plazo fijado para la entrega de los resultados de la investigación. Muchas ve-

ces este plazo es impuesto por un comitente que no tiene idea de la complejidad de la tarea. A veces la naturaleza del comitente, o más genéricamente el clima político-cultural, impone límites a la libertad de profundizar en una u otra dirección. Otras veces la naturaleza del problema investigado pone límites éticos a las técnicas que se podrían usar.

La selva por atravesar es recorrida por segmentos más o menos largos, en varias direcciones, de senderos ya trazados, más o menos transitados: son las técnicas que otros investigadores han ideado, modificado, desarrollado. Naturalmente, es mucho más cómodo recorrer senderos ya trazados, pero no se sabe si nos conducirán al claro buscado o a alguna otra parte.

La tarea del investigador/metodólogo consiste en escoger el camino, teniendo en cuenta la naturaleza de los senderos presentes en el terreno, el tiempo y los recursos disponibles (en particular, el grado de adiestramiento/predisposición de sus colaboradores para manejar las diversas técnicas). En este sentido, su tarea se asemeja a la combinación de factores productivos que el emprendedor efectúa en presencia de restricciones y limitaciones, disponiendo de recursos escasos (Schumpeter 1942) y de "economías externas".[8] También el emprendedor toma frecuentemente decisiones: en cada momento debe elegir si confiar una cierta tarea a colaboradores externos, o proceder por sí mismo pero siguiendo un sendero ya bien transitado, u otro sendero apenas trazado, o bien combinar varios senderos; decidirá también si recorrer fielmente estos senderos, o intentar pequeñas variaciones del recorrido (modificaciones a técnicas existentes), o bien internarse en la selva, imaginando procedimientos nuevos, a veces para confrontar los resultados con aquellos de las técnicas existentes.

Lo esencial del concepto de método está en lo siguiente: en la elección de las técnicas a aplicar, en la capacidad de modificar técnicas existentes, adaptándolas a los problemas propios, y de imaginar técnicas nuevas. Una vez que un procedimiento nuevo o una modificación de un procedimiento existente han sido ideados y son codificados y difundidos, éstos se reifican y se convierten en una técnica a disposición de la comunidad de investigadores: no ya en una capacidad privada sino en un objeto público. Sucede como con la *parole* de Saussure: una vez comunicada, ya es *langue*.

Los resultados de una investigación dependen del complejo de elecciones que el investigador/metodólogo ha realizado a lo largo de todo el recorrido. Por este motivo, en el informe de investigación se debería destinar amplio espacio para reportar y comentar las elecciones hechas, evaluando en lo posible sus impactos sobre los resultados. Lamentablemente, en cambio, es muy raro que ese espacio sea destinado (siendo Schussler 1982 una de las elogiables excepciones).

Como determinantes de esta situación —que demuestra una conciencia científica todavía embrional— convergen varios factores:

[8] Es decir, de recursos que la sociedad pone a su disposición (rutas, puentes, puertos, etcétera).

- el objetivismo del investigador, que cree debilitar el valor de sus resultados si admite cuánto dependen de las elecciones efectuadas (o que atribuye con gran fundamento tal convicción a su público);
- el hecho de que buena parte del público se comporta como si leyese una novela policíaca: quiere saber quién es el asesino y no se interesa por los detalles del proceso de investigación (lo lamentan Cohen y Nagel 1934: 399-400).
- En fin, el hecho de que los responsables de casas editoriales y revistas comparten en gran medida el objetivismo de autores y público, y el hambre de hechos del segundo, y deben tenerlos en cuenta aun si no los comparten.

Sucede entonces que incluso autores muy escrupulosos confinan el informe metodológico a un apéndice, imaginando que "aquel que quiera saber más de aquello que el investigador ha efectivamente hecho y del modo en que lo ha hecho [...] dispondrá de tanta paciencia como de curiosidad" (Calvi 1980: 21).

El hecho de que las técnicas no posean la naturaleza abstracta del método, sino que sean recursos tangibles y disponibles, no implica una menor nobleza o relevancia para el trabajo científico. La orientación de los grandes teóricos[9] que desprecian las técnicas es precientífica, dado que sólo a través de su empleo se puede de algún modo controlar empíricamente una teoría —e incluso una gran teoría. Además, sólo a través de un uso competente de las técnicas este control puede hacerse de modo confiable.

Simétricamente, cae en una postura errada aquel que fetichiza una o más técnicas, en el sentido de que desarrolla una gran pericia en su aplicación, pero las considera sólo como objetos *ready made*, carentes de historia y de profundidad: ignora y quiere ignorar el hecho de que cada técnica ha sido ideada y desarrollada en el marco de ciertos supuestos onto-gnoseo-epistemológicos,[10] que legitiman el recurso a ciertas manipulaciones de la información o a ciertas interpretaciones de los resultados.

Desde hace algún tiempo, esta tendencia a la especialización estrecha se ha manifestado —y ha sido relevada y denunciada en los Estados Unidos—: algunos metodólogos son en realidad especialistas en una o dos técnicas, y las aplican a todos los problemas que enfrentan sin cuidarse de su pertinencia. De esta forma producen resultados sin sentido,[11] y cuando se dan cuenta, no tienen soluciones, ya que no tienen una sabiduría y experiencia generalizada que les permitiría elegir las herramientas más adecuadas a su problema.

Se puede pensar que la superespecialización se desarrolla paralelamente a la tendencia al consumo de masas, porque ambas presentan dos características

[9] Los *grand theorists*, como los llamaba irónicamente Mills (1959), teniendo como blanco principal a Parsons.

[10] Sobre cómo es la realidad, cómo podemos conocerla y cuáles son los objetos y los límites de la ciencia.

[11] Como pasa cuando técnicas concebidas para variables cardinales se aplican a variables que no son ni ordinales.

estrechamente parecidas: la propensión a orientarse hacia el producto nuevo abandonando el antiguo cualquiera fuera su valor, y la propensión a orientarse hacia productos cada vez más artificiales y efectistas (que en nuestro caso quiere decir sobre todo *computer intensive*). Basta pensar en la rápida sucesión de modas que ha traído a la escena del análisis de datos: primero la regresión múltiple, luego el análisis factorial, luego el *path analysis*, luego los modelos log lineares, luego el análisis de las correspondencias, y así sucesivamente. Lo observaba de manera incisiva Labovitz:

> Que en el análisis de datos de las ciencias sociales imperan las modas es más que evidente. ¿Quién sigue usando el escalograma de Guttman? Esta técnica *demodée* ha sido sustituida por otras [...] Se usa la técnica en boga, sea apropiada o no. La gente se pregunta: "Quiero emplear el análisis factorial: ¿cuál es el problema?" O bien te dice: "Quiero utilizar el último grito de la moda, el *path analysis*: ¿Tienes algunos datos?" (1968: 221).

Se ha puesto como ejemplo el análisis de datos porque ha sido el teatro de este reciente desarrollo. El metodólogo italiano Ricolfi lo ha juzgado, un poco severamente, como un "progreso aparente" (1982: 338); también el norteamericano Davis piensa que "las técnicas estadísticas avanzadas no nos dicen mucho más que el análisis tabular de los tiempos de Lazarsfeld" (1987: 179). Pero, más allá del juicio sobre su utilidad, la opinión general es que las técnicas de análisis matemático-estadístico se encuentran actualmente mucho más desarrolladas que las técnicas de recolección de informaciones o de presentación gráfica de los resultados.

Hace algunos años otro metodólogo italiano lamentaba "el desequilibrio entre el desarrollo de estos métodos matemáticos y sus efectivas aplicaciones", sosteniendo que "el conocimiento de la matemática y de la estadística proporciona prestigio. Lo importante es que el investigador llegue a demostrar [...] que sabe dominar estos métodos" (Capecchi 1972: 39). "Los textos de metodología —ironiza Troy Duster— siguen afirmando que todas las técnicas tienen el mismo valor [...] De hecho, el programa de los cursos y el tipo de artículos aceptados para publicación pronto dan el mensaje a estudiantes y doctorandos: ¡las técnicas están estratificadas!" (1981: 112).

Naturalmente, podemos preguntarnos acerca del motivo de este mayor desarrollo y mayor prestigio de las técnicas de análisis de datos.[12] Se decía más arriba que la fascinación por las novedades es un motivo para la tendencia a la especialización; sin embargo, éste no es válido en lo que respecta a las técnicas de análisis en su conjunto: en el currículo de los departamentos de ciencias sociales norteamericanos y de Europa del norte las técnicas estadísticas no son novedad desde hace al menos treinta o cuarenta años.

[12] Van Meter, quien fuera presidente del comité de metodología de la asociación mundial de los sociólogos, ha afirmado tranquilamente, en una reseña dedicada por la revista de la UNESCO a los diversos sectores de la sociología: "El término 'método' significa habitualmente 'método de análisis de datos' a menos que se especifique otra cosa" (1994: 19).

Sin duda, vale en cambio el otro motivo, es decir, la propensión a procedimientos *computer-intensive* más bien que *field-intensive* (más similares a un juego electrónico que a un paseo con amigos por un bosque). Este factor actúa, sea por fuerza propia, sea por vía de la imitación de las ciencias físico-naturales, en las que las simulaciones computarizadas tienen un rol cada vez más importante: por esta razón parece destinado a ejercer una influencia duradera.

No se puede dejar de lado, en fin, una explicación de corte kuhniano: buena parte de los jóvenes emergentes en las ciencias humanas en los años sesenta y setenta han percibido la preparación estadística como el canal privilegiado para constituir rápidamente un patrimonio de *expertise* que sus mayores no poseían, lo que les permitía hacerse imprescindibles, y pretender a mediano plazo una sucesión anticipada (Collins 1975: 54).

3.6. Epistemología y gnoseología

En lo que respecta a la relación entre metodología (y método) y EPISTEMOLOGÍA, existe un consenso general sobre el hecho de que la metodología ocupa "la porción central de un *continuum* de análisis crítico [...] [entre] el análisis de los postulados epistemológicos que hacen posible el conocimiento social y [...] la elaboración de las técnicas de investigación" (Gallino 1978: 465). Ocuparse de metodología es encontrarse en una continua tensión dialéctica entre los polos de este *continuum*, porque "si la metodología abandona su lado epistemológico se reduce a una tecnología o una práctica que ya no controla intelectualmente. Si abandona el lado técnico, se transforma en una pura reflexión filosófica sobre las ciencias sociales, incapaz de incidir sobre las actividades de investigación" (Bruschi 1991: 41).

Resta agregar que en uno de los polos no se encuentra tan sólo la epistemología (reflexión sobre los fines, condiciones y límites del conocimiento científico) sino también la GNOSEOLOGÍA (reflexión sobre los fines, condiciones y límites del conocimiento *tout-court*). La gnoseología ha sido y es frecuentemente olvidada en las declaraciones como las mencionadas arriba. Las razones pueden ser dos, una lingüística y otra sustancial. El término gemelo (sería *gnoseology*) no existe en inglés, y por lo tanto nunca aparece en la imponente literatura anglonorteamericana sobre el tema. En inglés, de hecho, se usa *epistemology* para designar la filosofía del conocimiento (gnoseología) y en prevalencia *philosophy of science* por aquello que nosotros llamamos epistemología.

El segundo motivo puede derivar del hecho de que todos los intentos de dar un fundamento de certeza absoluta al conocimiento científico (los más recientes el fenomenismo de Schlick y del primer Carnap, y el fisicalismo de Neurath y del segundo Carnap) han fracasado frente a la constatación, inevitable en gnoseología, de que no puede haber nexos rígidos entre la esfera de los referentes (la realidad), la esfera del pensamiento y la esfera del lenguaje (Marradi 1994). Puede suceder que esta situación, aceptada con escaso entusiasmo por muchos de los que se ocupan de ciencia, haya provocado un sentimiento de

turbación frente a una disciplina que pone en evidencia los límites de las pretensiones cognoscitivas de las otras.

Si la metodología se sitúa donde se ha dicho, ¿donde se sitúa el método? En la sección previa se ha planteado que éste consiste esencialmente en el arte de elegir las técnicas más apropiadas para enfrentar un problema cognoscitivo, eventualmente combinándolas, comparándolas, aportando modificaciones e incluso proponiendo alguna solución nueva. Aquello que no surgió en las secciones precedentes y que quisiéramos subrayar antes de concluir este capítulo, es que el metodólogo no realiza estas elecciones sólo a la luz de sus competencias técnicas y de las experiencias de investigación propias y ajenas. Su propensión a utilizar una u otra técnica y su manera de interpretar las experiencias de investigación están condicionadas por sus opciones gnoseoepistemológicas: "Las soluciones técnicas presuponen soluciones metodológicas generales y estas últimas, por otra parte, implican que se den respuestas adecuadas a ciertas cuestiones epistemológicas" (Ammassari 1985: 178).

Capítulo 4

EL PAPEL DE LA TEORÍA EN LA INVESTIGACIÓN SOCIAL

La primera tarea del sociólogo debe ser por ello definir las cosas que trata, a fin de que se sepa —y lo sepa él también— cuál es el problema. (Durkheim 1895)

Un requisito básico de la investigación es que los conceptos, las variables, sean definidos con suficiente claridad para permitir que la investigación progrese. (Merton 1949)

4.1. ¿Qué es la teoría?

Proponer una definición del término 'teoría' conduce a una paradoja, ya que ésta sólo es definible dentro del marco de una teoría. Debido a que no existen definiciones extrateóricas en la investigación social, la propia teoría es alcanzada por esta restricción, que sitúa cada definición dentro de las categorías y el horizonte de una teoría dada.

Si bien toda definición es arbitraria y relativa al contexto de su elaboración, el problema en este caso particular es la coincidencia entre el concepto a definir y el contexto definitorio. Esto no significa que "teoría" sea un concepto indefinible sino que —dado que una teoría no puede validar la legitimidad de su propia definición— nos encontramos frente a un concepto polisémico. Existen múltiples definiciones de teoría cuya arbitrariedad sólo está limitada por el contexto teórico que las contiene, de tal modo que adherir a una definición supone ya la adhesión a una teoría previa.

Por otro lado, las utilizaciones del término no ayudan mucho para la definición de su significado. A veces hace referencia al pensamiento de un autor o de una corriente de pensamiento (la teoría marxista, la teoría psicoanalítica), otras alude a la obra de diversos autores relativa a una temática en particular (la teoría de la democracia, la teoría de la dependencia), también refleja contenidos metodológicos (la teoría general, las teorías de alcance medio) y otras veces se lo equipara al concepto de paradigma, en tanto una visión, cosmovisión o forma de comprender la realidad social. "Como muchas palabras excesivamente usadas, la palabra 'teoría' amenaza con quedar vacía de sentido" (Merton 1949/1964: 15).

Para la mayoría de los autores, uno de sus elementos definitorios es la interrelación de sus componentes. Éstos pueden estar expresados en forma de conceptos, categorías, hipótesis o como afirmaciones plausibles, pero lo relevante es que estén interconectados de alguna manera.

Merton es tal vez quien mejor expresa esta idea cuando afirma: "Se dice a veces que la teoría está formada por conceptos, afirmación que por ser incompleta no es verdadera ni falsa, sino vaga [...] Sólo cuando tales conceptos se relacionan entre sí en forma de un sistema, empieza a aparecer la teoría" (Merton 1949/1964: 99).

Desde diferentes perspectivas, los autores han compartido en sus definiciones de teoría este sentido relacional,[1] sin embargo, difieren respecto del proceso lógico que guía su generación (deductivo/ inductivo), de la forma de su evolución (acumulación/ ruptura) y de los roles que cumple en el proceso de investigación. Debido a que se le han asignado a la teoría características tan diversas respecto de su estructura lógica, su generación y su evolución, se vuelve difícil la identificación de sus elementos definitorios, aquellos que permitirían designar como "teoría" a un conjunto de enunciados.

El consenso respecto del carácter relacional de sus elementos no parece suficiente para su definición. El significado del concepto "teoría" se vincula también al papel que cumple en el proceso de investigación y a los usos que se le ha dado en las diferentes perspectivas. Aunque no siempre esté explícito, existe un consenso importante respecto de que la teoría abre un camino que conduce más allá de las observaciones, que apunta a develar relaciones entre los fenómenos que no son objeto de la experiencia inmediata. Como sostiene Weber: "La función específica de la ciencia es, a mi parecer, justamente la inversa: convertir en problema lo evidente por convención" (Weber 1917/1958: 233). Este proceso que consiste en poner de manifiesto lo oculto, dando nacimiento al descubrimiento científico, puede adoptar tanto la forma de la explicación como de la interpretación, pero siempre opera como elemento constitutivo de toda teoría científica. En palabras de Stinchcombe (1968/1970: 14): "[las teorías] son necesarias para explicar la pauta subyacente en las observaciones sobre el mundo".

4.2. El diálogo entre la teoría y los datos

La relación entre la teoría y los datos alude, desde diferentes perspectivas epistemológicas, a los objetivos del proceso mismo de investigación. En algunos casos, éste se pone en marcha con el fin de contrastar una teoría con los hechos, en otros con la intención de generar nueva teoría en la interacción con la realidad. En el primer caso, los datos se construyen a partir de la teoría; en el segundo, ésta se genera desde aquéllos.

Los análisis de la teorización refieren a diferentes tipos de razonamiento. Uno de ellos es la DEDUCCIÓN, basada en un proceso lógico mediante el cual se infieren conclusiones a partir de algunas premisas. Se trata de un procedimiento analítico donde, debido a las relaciones de implicación entre proposiciones de diferente nivel, la verdad de las conclusiones depende de la verdad de las premisas. La deducción es estrictamente un método de demostración. Aristóteles fue el primero que formuló de manera explícita los pasos de la inferencia deductiva sistematizándola en el *Organon*, y el método encontró su expresión más acabada en la geometría euclidiana.

La principal crítica a la deducción se orienta a su esterilidad como generadora de nuevos conocimientos, en la medida en que la verdad de las conclusiones ya está contenida en las premisas. Su carácter analítico la convierte en un camino válido para la demostración pero no para el descubrimiento. Los deductivistas modernos introducen elementos del empirismo en el método deductivo estableciendo la verdad o falsedad de las premisas a través de la confrontación empírica de las conclusiones, invirtiendo de este modo la inferencia. En este caso, la verdad de las premisas pasa a depender de la verdad de las conclusiones.

La verificación empírica del sistema hipotético-deductivo consiste en poner a prueba las hipótesis de nivel más bajo, cuya confirmación o refutación es el criterio por el cual se pone a prueba la verdad de todas las hipótesis del sistema. De modo que el establecimiento de un sistema como conjunto de proposiciones verdaderas depende de la aceptación de sus hipótesis de más bajo nivel. Las conclusiones son puestas a prueba a través de dos caminos: comparando sus conexiones lógicas con otras proposiciones (prueba de la coherencia interna de la teoría) y confrontándolas con los hechos (contrastación empírica).

Los defensores del método deductivo de comprobación han definido a la teoría científica a partir de características lógico-deductivas.[2] Karl Popper

[1] "La frase 'teoría sociológica' se refiere a conceptos lógicamente interconectados y de alcance limitado y modesto más bien que amplios y grandiosos" (Merton 1949/1964: 15). La tarea del teórico consiste en "formular hipótesis específicas, relacionadas entre sí, dando formulaciones nuevas a generalizaciones empíricas a la luz de las orientaciones generales" (*ibid.*: 98).

"La teoría consiste en relaciones *plausibles* propuestas *entre conceptos y conjuntos de conceptos*" (Strauss y Corbin 1994: 278).

"Los elementos de la teoría generados por el análisis comparativo son, primero, categorías conceptuales y sus propiedades conceptuales, y segundo, hipótesis o relaciones generalizadas entre las categorías y sus propiedades" (Glaser y Strauss 1967: 35).

[2] "Un sistema científico consta de un conjunto de hipótesis que forman un sistema deductivo; esto es, las hipótesis están dispuestas de tal manera que a partir de algunas de ellas, tomadas como premisas, se siguen lógicamente todas las demás hipótesis" (Braithwaite 1953/1960).

"Definimos una teoría como un conjunto de hipótesis estructurado por la relación de implicación o deducibilidad" (Galtung 1966).

"Una teoría de un fenómeno consiste en una serie de proposiciones, cada una de las cuales establece una relación entre propiedades de la naturaleza. Para constituir una teoría las proposiciones deben tener la forma de un modelo deductivo" (Homans 1970).

"Las teorías no consisten enteramente en esquemas conceptuales o tipologías, sino que deben contener proposiciones tipo leyes que interrelacionan dos o más conceptos o variables a un tiempo. Idealmente se debería alcanzar un sistema teórico deductivo completamente cerrado en el cual un mínimo conjunto de proposiciones se tomen como axiomas, de los cuales todas las otras proposiciones puedan deducirse por razonamientos puramente matemáticos o lógicos" (Blalock 1969).

"Las teorías se ponen a prueba comparando sus derivaciones empíricas con las observaciones" (Stinchcombe 1968/1970).

(1934) —quien formula con precisión el método deductivo de comprobación— sostiene que no existe justificación lógica que permita inferir enunciados generales de enunciados particulares: no existen argumentos para la inferencia inductiva. Sin embargo, dado que el carácter analítico del método deductivo no conduce al descubrimiento, la inducción aparecería como el camino válido hacia el conocimiento.

Si la deducción va de lo general a lo particular, la INDUCCIÓN recorre el camino inverso: a partir de la observación y la experimentación de hechos particulares se enuncian proposiciones universales que constituyen las leyes científicas, de tal modo que mediante la inducción se generaliza la experiencia. Francis Bacon —uno de los fundadores de la ciencia empírica— basó el método científico en la experiencia en el *Novum Organum Scientiarum*, y su expresión más desarrollada fueron las leyes de la mecánica. Para algunos defensores del inductivismo, como Reichenbach, la inferencia inductiva es de índole probable, en la medida que, si bien no permite afirmar con certeza la verdad de los enunciados generales, permite afirmar la mayor o la menor probabilidad de su verdad o falsedad; desde esta perspectiva los enunciados científicos resultan ser probabilísticos.

Dentro de la perspectiva inductiva, los datos adquieren un protagonismo diferente, dejan de ser la instancia de la comprobación teórica para convertirse en la fuente del descubrimiento teórico; la teoría resulta generada desde los datos (Strauss y Corbin 1994: 273). Los autores de la denominada *grounded theory*[3] (teoría fundamentada) sostienen que la investigación social y la generación de teoría son parte de un mismo proceso (Glaser y Strauss 1967, Glaser 1978: 273). Desde esta aproximación, la construcción de teoría supone una cercanía, un ajuste con los datos y una relación con los actores, necesarios para la aplicabilidad de la teoría resultante en la vida cotidiana, a diferencia del trabajo de investigación basado en materiales más remotos. Los datos no se consideran solamente como el resultado de las observaciones, sino como una construcción resultado de un proceso de interpretación de interpretaciones (Geertz 1973).

4.3. La teoría y su relación con el método

Uno de los desafíos en la enseñanza de las ciencias sociales es transmitir a los estudiantes que la ciencia es constitutivamente metódica, es decir, que sólo es posible conocer científicamente a través de algún método. Los alumnos, en general, se interesan poco por la metodología y se orientan más hacia la teoría, considerando a ambas como independientes y autónomas. Probablemente, una de las grandes responsabilidades por este error provenga de la *curricula* universitaria donde, generalmente, las materias son definidas como "metodológicas" o "teóricas" dejando a cargo de los estudiantes la integración entre ambas.

Cuando se logra algún éxito en la motivación estudiantil por la metodología, el segundo desafío consiste en transmitir que el método no es una receta

[3] Sobre *grounded theory*, véanse los capítulos 12 y 17.

estática, que ni siquiera es único y que está constituido por un conjunto de enfoques, procedimientos y herramientas sujetos a un continuo proceso de adecuación a los objetivos de investigación dentro de alguna perspectiva teórica. El método, en cada investigación, es el resultado de un diseño específico que se elabora, básicamente, a partir de dos elementos: los conceptos que provee la teoría y las características de los datos disponibles. Como la adecuación entre el andamiaje conceptual y los datos construidos por el investigador constituye un fenómeno único, en cada investigación particular se produce una recreación metodológica, resultado de la aplicación de los criterios de adecuación a los objetivos de investigación y a la información recolectada.

Nuestra intención es mostrar que teoría y método conforman una unidad constitutiva del quehacer científico, donde la primera establece el marco conceptual dentro del que se desarrolla el segundo, y éste fija el horizonte de aplicabilidad de la primera, en el marco del desarrollo técnico y tecnológico alcanzado.

Conocer científicamente constituye siempre un desafío; el camino de la ciencia es siempre riesgo y aventura para quien decida alejarse de los senderos que conducen a reconfirmar lo ya dicho. En este sentido, los pasos del procedimiento científico no pueden estar estrictamente establecidos de antemano. Los manuales presentan modelos ideales-normativos, generalmente estructurados por las etapas que debe seguir una investigación en un orden sucesivo, sin embargo, éstas no se cumplen siempre del mismo modo ni en el mismo orden en la práctica de la actividad científica (véase capítulo 5).

La particular relación entre teoría, método y datos le otorga a cada investigación un carácter específico. En la realidad del proceso de investigación cada investigador es a la vez un teórico y un metodólogo que adecua las técnicas y las herramientas con que cuenta (o las nuevas que desarrolla) a fin de alcanzar las metas que se propone (véase capítulo 3). En esta tarea, la íntima relación entre método y teoría define los límites de la actividad del conocer, capturada dentro del aparato conceptual teorético y condicionada en sus formulaciones por las potencialidades del método.

La teoría contiene tres elementos fundamentales que constituyen el marco o el horizonte para el desarrollo metodológico:

a) Las CATEGORÍAS, que operan como códigos interpretativos de la realidad y orientan la mirada hacia determinados aspectos y problemas. Los códigos permiten comprender la realidad y estatuirle un orden, una lógica interpretativa que le otorga sentido; constituyen las claves para la comprensión de esa realidad, sin ellos la asignación de sentido se vuelve una quimera, ya que diferentes sistemas de códigos llevan a resultados disímiles.[4] Debido a que miramos el mundo a través de un sistema de categorías, el método, para ser efectivo, debe construirse dentro de

[4] Un personaje de Eco expresa el poder revelador de los códigos: "Agliè me había enseñado que el secreto de las pirámides no se descubre calculándolas en metros, sino en antiguos codos" (1989: 17).

esos códigos, constituyéndose en la clave que abre el puente entre la teoría y la realidad interpretada. La elección de uno u otro sistema de categorías, de una u otra teoría, se fundamenta en la credibilidad de ésta dentro de la cosmovisión del investigador y en competencia con otras cosmovisiones. Según Weber, "las 'cosmovisiones' jamás pueden ser producto de un avance en el saber empírico y [...] por lo tanto, los ideales supremos que nos mueven con la máxima fuerza se abren camino, en todas las épocas, sólo en la lucha con otros ideales, los cuales son tan sagrados para otras personas como para nosotros los nuestros" (1904/1958: 46).

b) Los CRITERIOS DE RELEVANCIA, que establecen los límites del universo de problemas y de datos, dando respuesta a las preguntas: ¿cuáles son los problemas a atender?, ¿cuáles son las informaciones a tomar en cuenta en principio? La teoría contiene los criterios de relevancia y el investigador transforma la información en datos de acuerdo con esos cánones. El proceso de construcción de los datos se orienta por las pautas que la teoría establece, en este sentido, lo que constituye un dato para una investigación no lo es necesariamente para otra.

c) El registro de los límites del conocimiento científico en cada momento histórico, que fijan el punto de partida de la mirada y el horizonte cognitivo para la producción de nuevas teorías con mayor poder heurístico. La teoría expresa los límites del conocimiento adquirido, la única realidad conocida científicamente es la que está dicha en la teoría. No refleja especularmente la realidad como sostenían los positivistas, por el contrario, ésta es interpretada, seleccionada y comprendida desde la estructura de sentido que la teoría provee. Estructura que se asienta en actitudes valorativas, "[...] sin las ideas de valor del investigador no existiría ningún principio de selección del material ni conocimiento provisto de sentido de lo real en cuanto individual [...]" (*ibid.*: 71).

4.4. La teoría como aproximación a la verdad

Embarcarse en la aventura de la ciencia es sólo para espíritus agnósticos, para aquellos que se han resignado al carácter inalcanzable e inasible de la verdad, los que han superado el anhelo de lo absoluto y de las certezas. El camino de la ciencia no tiene un punto de llegada y los hallazgos de hoy podrán ser desechados mañana. Sin embargo, junto al carácter efímero del descubrimiento, las teorías y los métodos lo sobreviven a través de las actividades de relectura, reinterpretación, elaboración de hipótesis *ad hoc*, reformulaciones y síntesis.

Kuhn cuestiona la idea de la teoría como reflejo de la verdad y propone el carácter convencional de la base empírica y la competencia entre teorías rivales como proceso de avance de la ciencia:

> "Ningún proceso descubierto hasta ahora por el estudio histórico del desarrollo científico se parece en nada al estereotipo metodológico de la demostración de falsedad, por medio de la comparación directa con la naturaleza [...] El acto de juicio que conduce a los científicos a rechazar una teoría aceptada previamente, se basa siempre en más de una comparación de dicha teoría con el mundo. La decisión de rechazar un paradigma es siempre, simultáneamente, la decisión de aceptar otro, y el juicio que conduce a esa decisión involucra la comparación de ambos paradigmas con la naturaleza y la comparación entre ellos" (1972/1993: 128-129).

Existen, al menos, dos opciones para pensar la relación entre ciencia y verdad: una es considerar a la primera como un camino hacia la certeza, la otra es concebirla delineando senderos orientadores entre la incertidumbre. Ambas concepciones se basan én supuestos ontológicos y epistemológicos diversos. Los supuestos ontológicos se vinculan a la forma de concebir la naturaleza de la realidad: como un mundo autónomo cuya verdad es alcanzable, como un mundo sólo aprehensible a través de aproximaciones o como una realidad múltiple construida socialmente (véanse capítulos 2 y 3). Los supuestos epistemológicos expresan la relación entre el investigador que conoce y el objeto conocido, pensada por un lado como un descubrir y por otro como un construir; como una relación de independencia que sustenta la objetividad o como una relación que se construye en la interacción sujeto-objeto que da lugar a los resultados (Guba y Lincoln 1994).

A partir del neopositivismo se desvirtúa la relación especular entre teoría y realidad. La teoría ya no refleja ni expresa la verdad de los hechos; su éxito se basa en su credibilidad (la evaluación de otros); la verdad va a ser cada vez más una verdad por convención. En este espacio reside el carácter político de la teoría, no solamente como lo sostiene Max Weber por la participación de las ideas de valor del investigador al recortar infinitas causaciones según significaciones culturales, sino también al participar las ideas de valor de "los otros" que avalan o rechazan su legitimidad. En consecuencia, las teorías no compiten por su valor de verdad —en tanto adecuación objetiva— sino por la legitimidad de sus conclusiones, sus consecuencias, sus resultados y las estrategias de acción que de ellos se derivan.

4.5. Los usos de la teoría

> *Una ciencia empírica no puede enseñar a nadie qué* debe *hacer, sino únicamente qué* puede *hacer y, en ciertas circunstancias, qué* quiere.
> (WEBER 1904)

Desde diferentes perspectivas se han asignado a la teoría funciones o papeles diversos. Algunos se sitúan dentro del proceso mismo de investigación, mientras que otros lo trascienden proyectando las funciones de la teoría hacia la sociedad. Según la mirada, esta proyección puede basarse en su utilidad social o en su carácter crítico. Para los autores de la *grounded theory,* "la teoría debe tener al menos alguna aplicación práctica para otros colegas o grupos especí-

ficos bajo estudio. También puede ser útil para los tomadores de decisiones" (Strauss y Corbin 1994). Por su parte, los autores de la Escuela Crítica, desde una perspectiva marxista, pensaban que la ciencia debería ser un instrumento de transformación.

En relación con la función que cumple la teoría dentro de la investigación, un punto de divergencia estratégico entre deductivistas e inductivistas se relaciona con el lugar que ocupa en ese proceso. Es decir, si está ubicada al inicio y constituye un punto de partida o, por el contrario, si se va configurando a lo largo del proceso de investigación conformando una meta a alcanzar. Los primeros ponen el énfasis en la puesta a prueba, los segundos en el descubrimiento.

Para los teóricos LÓGICO-DEDUCTIVOS, la teoría constituye un modelo a contrastar —cuya elaboración es previa a la recolección de datos— que delimita el campo de la investigación, elabora el aparato conceptual, demarca la relevancia y focaliza la atención en problemas determinados. De este modo, la teoría traza un horizonte cognitivo demarcando un ámbito de registro a partir de un sistema conceptual-categorial, en función de un conjunto de criterios. El aparato conceptual y los criterios de relevancia operan orientando la atención del investigador hacia ciertos problemas, actores sociales y áreas de investigación. Los conceptos centrales de la teoría estructuran los ejes para la percepción de la realidad y cualquier nuevo concepto que se construya estará lógicamente articulado con ese cuerpo central. Los criterios de relevancia marcan el límite entre el interior y el afuera de la teoría, definen qué cosas la teoría comprende y cuáles no, qué aspecto de la realidad será tomado en cuenta y cuál será desestimado por irrelevante.

Los INDUCTIVISTAS, por el contrario, sitúan a la teoría al final del proceso investigativo; ésta surge desde el diálogo con los datos, constituye una meta a alcanzar, un producto a ser creado más que un modelo contrastable. El marco teórico, en este caso, se limita a ser un marco orientativo dentro del cual se elaboran las categorías en la interacción con los datos y se establece la relevancia a partir de lo observado.

Además de las relaciones lógicas entre sus elementos y de la facultad para develar las pautas subyacentes o emergentes, la teoría concierta un amplio consenso como guía para la investigación. Dentro del modelo hipotético-deductivo se trata de una guía estructurada, donde las categorías, conceptos y criterios de pertinencia son establecidos *a priori*. Dentro del modelo inductivo cualitativo su función es más elusiva, se trata básicamente de orientar en la incertidumbre.

La capacidad explicativa de la teoría es otro de sus rasgos recurrentes; en palabras de Stinchcombe: "La teoría debe crear la capacidad de inventar explicaciones" (1968/1970: 11). La capacidad de explicar —ya sea a través de enunciados generales o de entender interpretando, como sostiene Max Weber (1904/1958: 61)— está íntimamente vinculada al descubrimiento, a la búsqueda de hacer manifiesto lo velado. Ésta es, tal vez, la principal tarea de la teoría, mostrar lo no visto para dar cuenta de lo observado. Dentro de las perspectivas generalizadoras, herederas del positivismo, la EXPLICACIÓN entendida como un enunciado universal se asocia a la PREDICCIÓN. La teoría suma, entonces, a su capacidad explicativa, la predictiva, a partir de asociaciones que trascienden el

campo del resultado aplicándose a nuevas situaciones (Merton 1949, Glaser y Strauss 1967).

Weber incorporó dentro de la teoría las IDEAS DE VALOR del investigador, desde donde éste construye los criterios de relevancia que van a guiar todas las selecciones y decisiones que encauzan la actividad investigativa. De este modo, a través de la teoría se cuela la subjetividad en el proceso científico y, desde allí, la carga valorativa va a teñir todo el proceso de investigación, cada vez que se aplique un criterio inevitablemente preñado de significación cultural. La realidad es siempre abordada desde lo teórico, y lo teórico implica valores, presupuestos, visiones, etcétera.

> "El único resultado de cualquier intento de conocer la realidad 'sin presupuestos' sería un caos de 'juicios de existencia' acerca de innumerables percepciones particulares [...] Lo único que introduce orden en este caos es la circunstancia de que, en cada caso, sólo una *parte* de la realidad individual reviste para nosotros interés y *significación*, porque únicamente ella muestra relación con las *ideas de valor culturales* con las cuales abordamos la realidad" (Weber 1904/1958: 68).

Con la incorporación de la decisión valorativa, Weber inserta la política en el campo de la teoría. Sheldon Wolin (1960) propone reconsiderar su obra dentro de una concepción política de la teoría y de la actividad teórica. Según este autor, el teorizar tiene una significación política que trasciende a los propios teóricos.

Capítulo 5

EL DISEÑO DE LA INVESTIGACIÓN

5.1. Introducción

En este capítulo pretendemos realizar una introducción al diseño de la investigación. Para esto, resulta fundamental comenzar por la definición de algunas palabras clave que permitirán conferirle un sentido determinado al desarrollo del tema. La definición de éstas no es una cuestión sencilla. Por un lado nos enfrentamos al problema de que las ciencias sociales poseen muchos menos términos que conceptos, y entonces ocurre a menudo que a un mismo término se lo utiliza para referirse a cosas diferentes. Por otro lado, las definiciones conceptuales están vinculadas a tradiciones científicas, a distintas formas de entender la ciencia y la realidad. Por lo tanto, una definición implica una toma de posición.

En el contexto de este capítulo, vamos a utilizar el término INVESTIGACIÓN en un sentido relativamente amplio. A través de una primera aproximación enciclopédica, se puede presentar a la investigación como un proceso sistemático y organizado por medio del cual se busca descubrir, interpretar o revisar ciertos hechos, y cuyo producto es un mayor conocimiento de éstos.

Este proceso involucra una gran cantidad de decisiones y acciones articuladas y con distintos niveles de complejidad. Desde un punto de vista restrictivo —en el marco de la concepción tradicional de la ciencia (véanse capítulos 1 y 2)— se considera en particular al conjunto de decisiones y actividades orientadas a establecer relaciones causales entre distintos aspectos de la realidad bajo estudio o, al menos, al control empírico (e impersonal) de hipótesis, es decir, conjeturas acerca de la realidad que postulan cierto tipo de relación entre determinados aspectos de ésta. Preferimos no adoptar este sentido limitado de la expresión 'investigación científica', no sólo debido a diferencias con la postura subyacente sobre la ciencia en general que éste implica, sino también por la razón más instrumental de que una definición tan restrictiva nos enfrentaría a una serie de dificultades insalvables para el tipo de trabajo que realizamos con frecuencia en las ciencias sociales, y al que igualmente llamamos investigación. La definiremos, en cambio, como un proceso que involucra un conjunto de decisiones y prácticas (que a su vez conllevan la puesta en juego de instrumentos

conceptuales y operativos) por las cuales conocemos —lo que puede significar describir, analizar, explicar, comprender o interpretar— algunas situaciones de interés cuya definición y delimitación (o construcción) forma parte de las decisiones apenas aludidas. Debe quedar claro que estamos hablando de INVESTIGACIONES EMPÍRICAS, es decir, aquellas en las que se establece algún tipo de relación observacional con la situación de interés (o con algunos aspectos de ésta).

Esta relación observacional, sin embargo, no puede reducirse únicamente al empleo de una técnica de observación en particular. Como se ha dicho, ésta forma parte de un proceso complejo que excede los meros recursos instrumentales o técnicos: entre las muchas decisiones complejas que se toman en una investigación científica está justamente la de definir cuáles son las mejores herramientas técnicas (por ejemplo, de observación) de acuerdo con los objetivos cognitivos que se hayan planteado. Se insiste en este punto dado que en no pocas ocasiones el deseo de mostrar la actividad científica como algo totalmente controlado y controlable ha llevado a reducir todo el rico entramado de decisiones (muchas veces tácitas) implicadas en el proceso de conocimiento a una mera cuestión técnica —impersonal, objetiva— de control empírico de las hipótesis o incluso de análisis de datos. Recurriendo a una imagen propuesta por Marradi (1996), se puede sugerir que estas técnicas no son más que islas de conocimiento impersonal, formalizado y explicitado, que flotan en un mar de conocimiento tácito, no explicitado y difícilmente explicitable, característico de toda actividad humana, incluso de aquellas más formalizadas como la investigación científica.

Que el proceso de investigación esté atravesado por decisiones y operaciones que dependen en gran medida del conocimiento tácito no significa, sin embargo, que sea totalmente caótico, desordenado o imprevisible. Por el contrario, este proceso puede ser diseñado y planificado y, entre otras cosas, esto será un modo de hacerlo más controlable, en la medida que exigirá un esfuerzo por explicitar de modo anticipado las distintas decisiones involucradas en el mismo. Según las definiciones de diccionario,[1] un diseño es una "descripción o bosquejo" de algo, un "proyecto o plan", siendo el proyecto un "primer esquema [...] de cualquier trabajo que se hace a veces como prueba antes de darle forma definitiva", y el plan un "escrito en que sumariamente se precisan los detalles para realizar una obra".

Si bien estas definiciones de diccionario toman como intercambiables los términos 'diseño', 'proyecto' y 'plan' —al menos en algunos de sus sentidos— resulta importante señalar que, aunque indudablemente emparentados, en cuanto relativos a la investigación científica pueden tener connotaciones levemente distintas. Es frecuente limitar la noción de diseño al conjunto de decisiones teóricas y metodológicas que harán posible la investigación. Algunos autores, por ejemplo Samaja (2002a), usan la expresión proyecto de investigación en un sentido más amplio, para designar no sólo las decisiones metodológicas típicas del diseño, sino también aquellas cuestiones necesarias para el control de gestión de la investigación. En este sentido, el proyecto hace referencia a un documento que responde a aspectos contextuales e institucionales en el marco de los cuales se desarrollará la investigación: marco regulatorio, financiamiento, instancias de control institucional, etcétera. El plan de investigación, por su parte, se reserva en algunos casos para designar el esquema que explicita cómo se llevará adelante una investigación en particular desde el punto de vista operativo, e incluye por lo tanto la dimensión temporal (cronograma de actividades) y las relaciones de precedencia, simultaneidad, etcétera, entre las distintas acciones implicadas en el proceso de investigación.

[1] *Diccionario de la lengua española* de la Real Academia Española, vigésima segunda edición, 2001.

5.2. Características del diseño de la investigación

Habiendo hecho todas las consideraciones precedentes, estamos en condiciones de plantear los siguientes interrogantes: ¿cómo se diseña, cómo se planifica, cómo se piensa una investigación empírica?

La idea que muchos metodólogos comparten es que la investigación empieza en el mismo momento en que se la comienza a plantear o a concebir. Como se podrá constatar, el diseño de la investigación se asemeja a la idea que se tiene de diseño en otros ámbitos de la actividad humana, incluso en aquellos de la vida cotidiana con los que estamos más familiarizados. Una analogía muchas veces útil puede establecerse con el diseño de una casa: éste no es la casa en sí misma —como objeto tangible— sino una anticipación modélica y abstracta de ella, una representación esquemática de cómo será, y un detalle de las decisiones necesarias para poder construirla, convirtiéndola entonces sí en un objeto tangible. El grado de detalle de esta anticipación modélica puede variar desde un simple bosquejo —en el que se plasman algunas ideas generales orientadoras de aquello a lo que se aspira— hasta un conjunto de planos, dibujos de vistas, preselección de materiales, presupuestos, etcétera, que con gran nivel de detalle permitirán, por un lado, guiar el proceso de construcción efectiva y, por el otro, dar una idea clara de cómo será la obra terminada.

En la investigación empírica, esta anticipación en abstracto de las decisiones que se van a tomar durante su curso también puede variar; en este caso, tal variación se dará en función del grado de detalle y explicitación que las decisiones teóricas y metodológicas adquieran antes de establecer algún tipo de contacto observacional con los fenómenos de interés.

Esta concepción del diseño tiene dos consecuencias importantes. La primera es que por definición no es posible una actividad de investigación no diseñada; lo que pueden existir son actividades de investigación con distintos grados o niveles de diseño (entendiendo por esto al conjunto de decisiones previas al contacto observacional con los hechos/conductas/situaciones de interés —sea éste directo o mediado, naturalista o artificialmente recreado). La segunda es que, desde el punto de vista del grado de detalle del diseño de la investigación, no se puede pensar en tipos cerrados sino en un *continuum* de diseños posibles.

Este *continuum* está limitado por dos polos antagónicos que habitualmente se conocen como DISEÑO ESTRUCTURADO y DISEÑO EMERGENTE (Valles, 1997).

En el primer extremo todo está absolutamente planificado de antemano: nada de lo que se lleve a cabo durante el proceso de investigación exigirá decisiones no previstas o se apartará de aquello ya pensado. La investigación implicará solamente una aplicación fiel del plan estipulado. En el segundo extremo, en cambio, nada está planificado: las decisiones que harán posible la investigación irán "emergiendo" durante el proceso mismo, con una lógica de *feed-back* a partir de los dictados que surjan del trabajo observacional de campo.

Se prefiere conceptualizar estos polos como tipos ideales, dado que en la práctica no se puede concebir una investigación que —por más prolija y completamente planificada que esté— no exija decisiones posteriores mientras se va desarrollando, no presente situaciones imprevistas o demande del conocimiento personal y tácito de aquellos que la llevan adelante —en momentos insospechados— para que el proceso llegue a buen puerto. Por su parte, un modelo en el que todas las decisiones vayan emergiendo durante el proceso mismo de investigación tampoco es imaginable: siempre habrá al menos un mínimo de decisiones previas que hagan posible la investigación.

En conclusión, llevada a sus últimas consecuencias la idea de diseño estructurado no es realista; es sólo un reflejo de los múltiples intentos de reducir la práctica científica a la completa certidumbre, a lo explicitable, al mero conocimiento impersonal formalizado (véase apartado 3.2). La idea de diseño emergente, por su parte, es un contrasentido, una contradicción en los términos que implica la negación de la idea misma de diseño: ¿en qué sentido se puede hablar de algo diseñado si no ha sido diseñado? Cabe señalar al respecto que el diseño emergente, supuestamente "liberador" de las ataduras de concepciones rígidas heredadas del positivismo (que exigen estrictos diseños de contrastación de hipótesis), también ha sido fuertemente criticado porque tiende a promover un tipo de investigación que se presenta como ateórico, y en el cual todo surge de los datos, como si la realidad "hablara" por sí misma. Desde este punto de vista, en el plano epistemológico el diseño emergente representa un retroceso desde las perspectivas neo y pospositivistas hacia formas previas de la ciencia convencional, caracterizadas, entre otras cosas, por un inductivismo ingenuo.

En la práctica de la investigación, en cambio, se dan situaciones intermedias entre estos polos ideales, así como diferentes combinaciones de algunas de las características de cada uno de ellos. Lo que se plantea entonces es la idea de DISEÑOS FLEXIBLES, que pueden ser más o menos estructurados según —como se decía más arriba— el grado de detalle que adquiera la planificación previa. Se asume, sin embargo, que existe un conjunto de decisiones presentes en cualquier tipo de investigación; este "mínimo" de diseño implica decisiones que se pueden agrupar en cuatro grandes conjuntos:

- Decisiones relativas a la construcción del objeto / delimitación del problema a investigar.
- Decisiones relativas a la selección.
- Decisiones relativas a la recolección.
- Decisiones relativas al análisis.

Si bien todas ellas están presentes de algún modo u otro en el diseño de la investigación, adquirirán diferentes características en cada caso particular. Resulta útil para clarificar esta cuestión una distinción ya presente en la filosofía griega clásica: acto y potencia. A continuación se presentan dos ejemplos —uno relativo a las decisiones de selección y otro a las de recolección— que permitirán esclarecer este argumento:

1) La selección de los sujetos a encuestar en el marco de un sondeo de opinión se realiza habitualmente antes de comenzar el trabajo de campo —especialmente si se pretende tener cierto control sobre las posibilidades de generalización de los resultados obtenidos en la muestra. Obviamente, sin una muestra de sujetos a encuestar no es posible encuestarlos. En este caso, la selección es ya acto en la fase de diseño: la muestra es un producto tangible (por ejemplo, un listado de personas) y esto puede haber sido el resultado de complejos procedimientos de selección sin que haya mediado un acercamiento directo a la población de referencia. En otros casos, estas decisiones de selección están tan sólo vislumbradas, son mera potencia. Pensemos en un trabajo etnográfico, en el que las observaciones de campo serán complementadas con entrevistas en profundidad (véanse capítulos 10 y 12), porque esto ha sido considerado pertinente en relación con los objetivos cognitivos que se persiguen. Imaginemos incluso que se ha delineado una estrategia y un conjunto de criterios para seleccionar a los entrevistados. Sin embargo, la materialización del sujeto a entrevistar muy probablemente no se logre hasta tanto no se adquiera cierto conocimiento directo, a través de la observación de campo, de las personas que conforman una situación social específica. En otras ocasiones, los resultados de dicha observación podrían cambiar el panorama previamente vislumbrado, haciendo necesaria una revisión de los criterios previos sobre quiénes entrevistar, teniendo el cuenta el lugar que las distintas personas ocupan en la situación de interés. Es decir, por ejemplo, que alguien que de antemano se pensaba clave como potencial informante se descubre irrelevante, y alguien que no había sido siquiera tenido en cuenta se revela un informante excepcional.

2) Para realizar el sondeo al que se viene haciendo referencia, se recurrirá a un cuestionario estandarizado que ha sido cuidadosamente diseñado de antemano (incluso puesto a prueba para controlar su efectividad). Si bien estas pruebas previas implican en general algún tipo de contacto directo con sujetos análogos a aquellos a quienes está destinado, el instrumento de recolección de informaciones —el cuestionario— es un producto concreto que ha sido diseñado íntegramente antes de llevar a cabo el trabajo de campo. La recolección se limitará a aplicar de un modo uniforme el mismo cuestionario a todas las unidades de la muestra, siguiendo idéntico orden y utilizando incluso las mismas palabras, que son las que han sido escogidas en la redacción de las preguntas. Aquí, como se puede constatar fácilmente, muchas de las decisiones de recolección, especialmente las relativas al instrumento que la hará posible,

son acto como producto del diseño: sin cuestionario no es posible realizar las encuestas. En cambio, pasando ahora al ejemplo de las entrevistas en profundidad, el contenido y orden específico de las preguntas no termina de develarse hasta tanto éstas no ha sido realizadas. La entrevista es una actividad diseñada, y entre otras cosas esto se manifiesta en que se construye una guía orientadora, que será más específica y detallada cuanto más se sepa del tema y de los sujetos a entrevistar. Pero la guía, aun en los casos de máximo nivel de detalle posible, no incluye todas las cuestiones que probablemente emerjan durante la entrevista ni prescribe un orden determinado para hacerlas. La riqueza de la entrevista está en que, en cuanto situación de interacción social, permite acceder a cierto tipo de informaciones para las cuales no se contaba con un instrumento previo completamente articulado. De haberlo, dejaría de ser una entrevista en profundidad para pasar a ser una encuesta. En este caso, entonces, nos encontramos con una combinación variable de decisiones que son acto en el diseño y otras que tan sólo son potencia, cobran acto al realizar la entrevista.

Las decisiones de selección (de personas, de pueblos, de espacios, de momentos, de documentos, etcétera), de recolección (¿por qué medios se obtendrá la información necesaria a los fines de la investigación?) y de análisis (¿qué técnicas y herramientas serán empleadas para ordenar, resumir, dar sentido a la información recolectada?) dependen del problema que se aborde. Por lo tanto, antes que nada, los diseños de investigación incluyen cuestiones relativas a la delimitación del problema de interés que —como se acaba de sugerir— condicionará el resto de las decisiones, especialmente en la medida en que todas ellas deberán ser instrumentales al logro de los objetivos cognitivos que el problema planteado conlleva.

En general, las investigaciones comúnmente llamadas cuantitativas (aquí estándar) requieren de diseños más estructurados: por sus características y por su naturaleza —tal como hemos visto en los ejemplos— exigen contar de antemano no sólo con una reflexión genérica acerca de qué se necesitará hacer, desde el punto de vista técnico-metodológico, sino con instrumentos concretos sin los cuales el trabajo observacional de campo no resultaría posible. En efecto, al momento de la recolección de información ya se debe contar con un equipo de trabajo entrenado, con una muestra, con un cuestionario estructurado y estandarizado, todo esto resultado de una planificación meticulosa. Desde un punto de vista técnico, entonces, este diseño más estructurado no sólo es posible, sino también necesario, cuando se piensa en una investigación como la apenas aludida, en la que se pueden "aislar" analítica y operativamente los distintos aspectos del proceso de investigación, y se puede incluso dividir el trabajo vinculando cada una de las tareas con personas y grupos relativamente independientes. Por ejemplo, en el caso del sondeo, parte del trabajo —la recolección de informaciones— es realizada en general por un grupo de personas que no ha participado activamente en el diseño del cuestionario y que tampoco se involucrará en el análisis estadístico de sus resultados.

Las investigaciones comúnmente llamadas cualitativas (no-estándar) se prestan habitualmente a diseños más flexibles: hay cuestiones que se pueden definir de antemano, pero hay muchas otras que no pueden ser definidas con anticipación y que deberán ser decididas a lo largo del proceso de investigación y en función del acercamiento a los objetos o sujetos de interés. Esto es así porque hay procesos, detalles, dimensiones fundamentales para la investigación que sólo pueden descubrirse mientras se observa directamente a los sujetos en sus espacios cotidianos, o cuando se entabla un diálogo con ellos. En estos casos se va a requerir de un mínimo de decisiones previas de diseño que orienten el estudio (y que muy fundamentalmente, entre otras cosas, justifiquen la opción por este tipo de recorrido investigativo), pero también habrá otro tipo de decisiones que se irán tomando mientras se desarrolla la investigación. Maxwell (1996) ha propuesto en este sentido el concepto de DISEÑO INTERACTIVO: un modelo holista y reflexivo de investigación en el que sus diferentes instancias se relacionan y afectan mutuamente sin seguir una lógica secuencial. En efecto, en estas investigaciones no resulta fácil separar los distintos aspectos del proceso: selección, recolección y análisis, al menos en algunos sentidos o aspectos, suelen tener una relativa simultaneidad. Tampoco es habitual operar una tajante división del trabajo: es el mismo equipo, o incluso el investigador en soledad, quien diseña y ejecuta, selecciona, recolecta y analiza.

5.3. La elección de un tema, la delimitación del problema y su operativización[2]

Como se indicó más arriba, la primera y fundamental cuestión en un diseño de investigación —que va a orientar y condicionar el resto de las decisiones— es la delimitación de un PROBLEMA DE INVESTIGACIÓN: ¿qué es exactamente lo que se desea conocer, y por lo tanto investigar?

Valles (1997: 83) indica que la formulación del problema es "un proceso de elaboración que va desde la idea (propia o ajena) inicial de investigar sobre algo, hasta la conversión de dicha idea en un *problema investigable*". Estas ideas iniciales pueden concebirse como TEMAS DE INVESTIGACIÓN —mucho más generales que los problemas, y no directamente abordables— que se relacionan con determinados "recortes" de la realidad propios de la estructura y del estado de desarrollo de la disciplina científica dentro de la que se inscribe el investigador, y/o de lo que habitualmente se conoce como "agenda de investigación" (cuestiones reconocidas como prioritarias en un lugar y momento específicos).

Se suele identificar como fuentes de temas (y consecuentemente de problemas de investigación) las sugerencias de profesores e investigadores experimentados, las convocatorias institucionales para acceder a becas o para finan-

[2] Se usa el neologismo 'operativización' en un sentido amplio —y en lugar del más difundido 'operacionalización'— para referirse al proceso por el cual se hace operativamente investigable un problema de interés. Nótese que en la metodología este término se usa comúnmente en un sentido más restringido (véase apartado 7.1).

ciar proyectos, la lectura de la literatura científica y la experiencia personal (Valles, 1997). La elección de un tema es, por lo tanto, una elección condicionada. En un sentido general, se puede afirmar que todo proceso de investigación se da en el marco de un contexto —cultural, social, político, económico e institucional— que configura las condiciones históricas que lo hacen posible (Samaja 2002b). En un sentido más específico, y por más independientes que pretendamos ser en la elección de un tema, este proceso va a estar influido por la tradición de la disciplina en la que se inscribe,[3] por el modo y las características que haya adquirido la socialización científica del investigador (qué perspectivas, lecturas, autores, conceptos, etcétera, le son familiares y han moldeado su forma de entender la realidad) y por otros aspectos más mundanos, pero no menos decisivos, como por ejemplo el financiamiento selectivo dado por los organismos de gestión de la actividad científica a unos u otros temas de investigación, junto con los términos de referencia que el otorgamiento de tal financiamiento conlleva.

Ahora bien, estos grandes temas de interés de una disciplina en un momento determinado —como ya se indicó— no son directamente investigables, entre otras cosas por su grado de abstracción, complejidad y amplitud. Sin embargo, todo tema puede llegar a convertirse en un problema de investigación; es decir, puede dar lugar a una pregunta o a un conjunto articulado de preguntas específicas que se puede abordar e investigar empíricamente. Para esto será necesario identificar propiedades conceptuales que, al realizar la investigación, van a constituir el foco de la observación empírica y fijar límites temporales y espaciales (el ámbito, véase apartado 6.1) en el marco de los cuales serán analizadas las propiedades conceptuales en cuestión y sus relaciones.

Un problema, entonces, siempre se define y construye a partir de un tema más amplio. Esta definición implica fundamentalmente un movimiento progresivo desde lo abstracto y general del tema hacia lo concreto y específico del problema. Pero el problema de investigación no surge espontáneamente, de un momento a otro. El proceso de su construcción, en cambio, puede ser extremadamente complejo y no lineal, e implica la puesta en juego tanto de saberes tácitos como de la experiencia; no hay una técnica (en el sentido que se le ha dado a este término en el apartado 3.5) para la formulación de problemas de investigación. Metafóricamente, se podría pensar esta cuestión como un recorrido espiralado descendente en el que los anillos de la espiral se hacen cada vez más pequeños, representando de este modo la mayor focalización gradual que van adquiriendo los asuntos de interés, hasta llegar a un núcleo que constituye el problema de investigación.[4]

3 Nos referimos a las tradiciones como resultado de construcciones sociales sedimentadas; se trata en definitiva de formas de entender y abordar la realidad compartidas por (parte de) la comunidad científica de una disciplina en un momento y lugar dados, y que marcarán los límites dentro de los cuales es posible concebir un problema de investigación.

4 Sobre la base de estas consideraciones se puede afirmar que la cantidad de problemas que se puede plantear en el marco de un mismo tema es enorme, y probablemente cada uno de ellos exigirá para su resolución distintas estrategias metodológicas y herramientas técnicas. Nótese que

Este proceso es posible a partir de un conjunto de prácticas que se conocen como INDAGACIONES PRELIMINARES. Con esta expresión se suele aludir a las lecturas más bien generales que permiten ir conquistando familiaridad con un tema; pero también se podría pensar en otras modalidades de indagación, como por ejemplo entrevistas con referentes en la materia. En el proceso espiralado al que se ha hecho referencia, estas indagaciones preliminares, en la medida en que se gana conocimiento sobre una cuestión, van haciendo posible identificar aspectos problemáticos; delinear preguntas específicas que aparecen como relativamente inciertas y dignas de profundización a los ojos del investigador. Al plantearlas, se empieza a darle forma al problema de investigación. Pero las indagaciones en torno de éste no cesan, sólo que ahora, mucho más focalizadas, constituyen lo que se define como ESTADO DE LA CUESTIÓN.

En efecto, uno de los primeros desafíos que deberá enfrentar el investigador, una vez definido el problema, es el análisis de sus antecedentes. Como decíamos, no se trata ya de esas indagaciones preliminares fundamentales para su delimitación, sino de pesquisas más específicas relativas al estado de conocimiento sobre el problema en cuestión, una revisión de la literatura científica directamente relevante en función del problema planteado: ¿se han hecho investigaciones sobre esa problemática?, ¿de qué tipo?, ¿a qué conclusiones llegaron?, ¿qué instrumentos se utilizaron en ellas? Cabe señalar que el estado de la cuestión no se limita simplemente a una reseña sintética de estas investigaciones previas; es recomendable además establecer un "diálogo" crítico con ellas. Por otra parte, a la hora de elaborar un estado de la cuestión es muy importante mantener el foco, evitando dispersarse *ad infinitum* en el repaso de antecedentes que sólo están vinculados muy marginalmente con aquello que se investiga. Los criterios para determinar la relevancia de investigaciones precedentes son fundamentalmente dos: afinidad temática y afinidad contextual (de ámbito y de unidades). Cuanto más afín es el tema abordado y más similar el contexto de una investigación precedente, más relevante resulta como antecedente del problema en consideración.

La delimitación de un problema conlleva la formulación de los OBJETIVOS de la investigación. Entre el problema y los objetivos existe una relación lógica de mutua implicación; sólo que si los problemas se plantean en forma de interrogantes, los objetivos se expresan en cambio por medio de proposiciones. Éstas contienen los mismos conceptos teóricos fundamentales que dan sentido al

pequeños cambios en los términos que se usan en esta delimitación derivan en problemas de investigación completamente distintos. Por ejemplo: investigar las "estrategias didácticas que desarrollan las maestras de EGB en las escuelas rurales de la provincia de Buenos Aires en la enseñanza de la lengua" no es lo mismo que investigar las "estrategias didácticas que desarrollan las maestras del primer ciclo de EGB en las escuelas privadas de la capital de la provincia de Buenos Aires para la enseñanza de la lengua". ¿En qué radica la diferencia? A pesar de la aparente similitud, se trata de actores y espacios geográficos distintos, cuya variación se expresa a través de palabras como "rurales", "privadas", "provincia" y "ciudad", que hacen cambiar completamente el foco y la dirección de la investigación.

problema de investigación; pero a través de ellas "el investigador postula [la] intención, generalmente explicitada por medio de un verbo (analizar, explicar, comprender, describir, explorar, etcétera), de abordar un sector de la realidad en un espacio y tiempo determinados" (Sautu *et al.*, 2005: 36). Debe quedar claro, por lo tanto, que nos referimos primordialmente a objetivos cognitivos, es decir, a aquellos orientados a acrecentar el conocimiento sobre un fenómeno dado. Éstos no han de confundirse con los objetivos ligados a la intervención o resolución práctica de un problema social. Los objetivos de la investigación habitualmente se clasifican como generales y específicos, siendo estos últimos aquellos que se derivan lógicamente de los primeros y cuyo cumplimiento concurre (o es incluso instrumental) al logro de ellos.

Para decidir la estrategia metodológica y escoger los instrumentos adecuados a los objetivos resultará necesario desentrañar el problema de investigación y analizar todas sus consecuencias en términos de una posible respuesta empíricamente construida. En algunas ocasiones, una respuesta tentativa existe antes aún de establecer una relación observacional (empírica) con los aspectos de interés de la realidad social. Desde cierto punto de vista, se puede afirmar que siempre existirá algún tipo de supuesto —en sentido amplio— acerca de aquello que se investiga. Los supuestos son inherentes al planteo de las preguntas de la investigación, al menos en la medida que están implícitos en las definiciones y perspectivas teóricas desde las cuales se las construye. Sin embargo, estos supuestos no siempre adquieren un grado de articulación tal de poder plantearse como HIPÓTESIS, es decir, como conjeturas que postulan un cierto tipo de relación entre los aspectos observables de la realidad que el problema de investigación ha delimitado. Contra la idea más ampliamente difundida, Singleton *et al.* (1988/1993: 88-89) afirman —a nuestro juicio de manera correcta— que las hipótesis, entendidas en sentido estricto, bien pueden ser el resultado de una investigación más que su disparador inicial. En otras ocasiones, ellas no se hacen explícitas desde un principio e igualmente guían, de manera tácita, todas las actividades de investigación. Sin embargo, cuando el objetivo del estudio implique la contrastación de hipótesis, éstas deberán estar precisamente formuladas, especulando sobre la naturaleza y la forma de una relación. Los autores antes citados sostienen que los modos más comunes en que aparecen las hipótesis en el trabajo científico son cuatro:

a) declaraciones condicionales (por ejemplo: si se da el fenómeno *y* se dará también el fenómeno *x*);
b) funciones matemáticas, que representan la expresión más parsimoniosa —y con fuerte carácter predictivo— de una relación;
c) declaraciones continuas (por ejemplo: cuanto mayor sea *x* mayor será *y*), y
d) declaraciones diferenciales, que afirman que una variable difiere en términos de las categorías de otra variable.

Independientemente del grado de formalización de la hipótesis, coincidimos con Maxwell (1996) en otorgar al problema el lugar central y la función rectora en la investigación: podemos no tener respuestas previas claras sobre un fenómeno bajo estudio, pero no podemos no contar con preguntas que orienten su abordaje y hagan posible su delimitación.

Para hacer operativo un problema de investigación, se lo puede pensar como la explicitación del interés del investigador por conocer "algo" de "alguien". Ese "algo" o "qué" que se quiere conocer se expresa a través de las propiedades conceptuales (y sus relaciones) cristalizadas verbalmente en las preguntas que guían la investigación. Por otra parte, ese "algo" se refiere siempre a "alguien",[5] un "quién" que está temporal y espacialmente situado, y que técnicamente —en cuanto definición abstracta— se denomina UNIDAD DE ANÁLISIS.

Con respecto al "algo" de interés deberemos comenzar por su CONCEPTUALIZACIÓN. Se trata de dar definiciones precisas de qué se entiende en el contexto de la investigación por aquellos términos clave que expresan verbalmente el foco de nuestra atención, definiciones que en su conjunto conformarán un MARCO CONCEPTUAL. El estado de la cuestión —a lo que se hizo mención más arriba— podrá constituirse en fuente de referencia para tales definiciones, permitiendo identificar perspectivas teóricas empleadas en trabajos anteriores y ayudando a definir los términos clave en relación con usos previos de ellos en la disciplina. Pero es importante no confundir el estado de la cuestión con lo que tradicionalmente se denomina MARCO TEÓRICO, es decir, el "corpus de conceptos de diferentes niveles de abstracción articulados entre sí que orientan la forma de aprehender la realidad" (Sautu *et al* 2005: 34), y que en el nivel más concreto incluye al marco conceptual.

Habitualmente, las definiciones conceptuales son muy abstractas, y se deberá entonces, especialmente en el caso de investigaciones de tipo estándar (o cuantitativas), seleccionar, a través de un proceso de OPERATIVIZACIÓN, los indicadores o referentes empíricos de ese "algo" que ahora ya representamos a través de definiciones conceptuales precisas. En otras palabras, deberemos llevar al plano de lo observable ese "algo" abstracto, y luego elegir las herramientas adecuadas para observarlo. Estas operaciones se basarán en un conjunto de decisiones que hemos denominado "de recolección", y que trataremos con detalle en los capítulos 7, 8, 9, 10, 11, 12 y 13.

Pero también se deberá operativizar la unidad de análisis, el "alguien" de interés al que se refiere la investigación. En primer lugar, dando una definición conceptual clara. En la práctica esto implica decidir criterios de inclusión y exclusión, ya que toda unidad puede concebirse como especie de un género mayor. Los criterios de inclusión y exclusión serán categoriales, temporales y espaciales. Por ejemplo, si estamos investigando sobre adolescentes, una definición de la unidad de análisis podría determinar, entre otras cosas, los límites mínimo y máximo de edad cronológica de una persona como para ser considerada adolescente en el marco de la investigación, y esto dependerá obviamente

[5] Nótese que decimos "alguien" para referirnos a las unidades de interés, pero éstas no tienen que ser necesariamente personas, podrían ser instituciones, familias, pueblos, naciones, productos de la cultura, etcétera (véase apartado 6.1).

de opciones teóricas que se adopten con respecto a la adolescencia y al modo de conceptualizarla. Pero los criterios —como acabamos de decir— también incluirán cuestiones temporales y espaciales: nos interesa saber algo sobre el adolescente así definido que habita ¿qué espacio geográfico? y ¿en qué período de tiempo? Este conjunto de criterios categoriales, espaciales y temporales también permitirán identificar la población, es decir, el conjunto de todos los casos que se correspondan con la definición conceptual que se haya dado de la unidad de análisis. En la mayoría de las situaciones, la población de interés resulta inabarcable con los recursos disponibles para la investigación. Esta limitación exigirá seleccionar alguno (o algunos) de sus casos a los fines de observación e indagación. Por lo tanto, parte del diseño de la investigación consistirá en la construcción de una muestra, que según el tipo de estudio va a seguir distintos criterios y asumirá diferentes características. Si bien ésta no es la única operación de selección presente en una investigación, es sin dudas la más crucial, y se inserta en el marco de lo que hemos llamado genéricamente "decisiones de selección", que se serán tratadas en el capítulo 6.

Sólo cuando todas estas decisiones hayan sido tomadas resultará posible realizar el trabajo de campo,[6] que podría limitarse, por ejemplo, a la aplicación de un cuestionario estandarizado (véase capítulo 11) a la muestra seleccionada, o implicar otras acciones como la realización de entrevistas en profundidad (véase capítulo 12), observaciones participantes (véase capítulo 10), etcétera. Los resultados de estas prácticas de recolección de información (que será obviamente información relevante sobre los "alguien" y con respecto al "algo" de interés), deberán ser procesados, ordenados, sistematizados y analizados. Estas operaciones constituyen el núcleo de lo que hemos llamado "decisiones de análisis", y serán abordadas en los capítulos 15, 16 y 17.

5.4. Formalización del diseño: el proyecto de investigación

Como sostiene Samaja (2002a: 204), "el objetivo central de la fase formulativa [de la investigación] es construir *escrituralmente* el objeto [de ésta] en su versión global". Pero en tanto la ciencia es una actividad profesionalizada e institucionalizada, esta versión escrita y formal deberá seguir las pautas implicadas en el proceso orientado a su gestión y control. En este sentido, se puede afirmar que un diseño de investigación se vuelve proyecto en el marco de un contexto institucional específico, por ejemplo, cuando se siguen las pautas que una universidad fija para la presentación de un plan de tesis o las condiciones que una agencia estatal de investigación establece para acreditar y

[6] Cabe recordar algo que se ha planteado más arriba: la completitud de todas estas decisiones como prerrequisito del trabajo de campo (es decir decisiones que sean acto en el diseño) se da sólo en algunos tipos de investigaciones (especialmente las estándar). En muchas otras ocasiones, el trabajo de campo no depende de que absolutamente todas las decisiones de recolección y selección ya hayan sido tomadas en la fase de diseño, sino que por el contrario es el desarrollo de dicho trabajo el que hace posible que muchas de ellas se completen y adquieran su forma definitiva.

financiar un estudio. A continuación se presentan, tan sólo a título ilustrativo, algunos ejemplos de los formatos que distintas instituciones científicas, en diferentes instancias, exigen para la presentación escrita de un proyecto de investigación.

Dei (2006), aun reconociendo que cada institución suele tener una versión propia, propone un esquema que contiene todos los puntos que a su juicio exigen la mayoría de las universidades argentinas en lo que concierne a la presentación de un proyecto o plan de tesis. Su esquema, que considera acorde al proceso psicológico de ideación de la propuesta de una investigación orientada a la obtención de un grado académico determinado, incluye, en lo que respecta al contenido del plan, los siguientes apartados:

1. Fundamentación del tema elegido.
   1.1. Antecedentes del tema.
   1.2. Estado actual del tema.
   1.3. Planteo del problema.
   1.4. Alcances y limitaciones de la propuesta.
   1.5. Aportes teóricos y/o prácticos al campo disciplinar.
2. Objetivos y/o hipótesis de trabajo.
3. Metodología y plan de actividades.
4. Bibliografía.
5. Resumen.

Por su parte, el documento escrito que debe presentar un candidato a beca de posgrado tipo I del Consejo Nacional de Investigaciones Científicas y Técnicas de la Argentina (CONICET) incluye los siguientes puntos:[7]

- Título.
- *Objetivos*: Se espera que en este apartado el candidato presente el objetivo general y específico de su investigación. Según las propias definiciones de CONICET, "el objetivo específico refleja el resultado esperable en el plazo previsto para la realización del plan. El objetivo general, al cual contribuye el objetivo específico, se orienta hacia resultados posibles de obtener en el largo plazo en la línea de investigación mediante trabajos ulteriores no incluidos en el plan".
- *Antecedentes*: Se debe "indicar el marco teórico de la investigación y las hipótesis de trabajo propuestas consignando sobre qué otros trabajos de investigación propios o de contribuciones de terceros se basan".
- *Actividades y metodología*: En esta sección se deberán "enumerar las tareas a desarrollar y las metodologías [...] y técnicas a emplear".
- *Factibilidad*: Esto se refiere a la infraestructura, los servicios y el equipamiento con los que cuenta el lugar de trabajo en el que se piensa de-

[7] Fuente: <www.conicet.gov.ar/becas/instructivo_becas2006.php> (fecha de consulta: 8/11/2006).

sarrollar la investigación, los que deberán ser adecuados para que el plan de trabajo resulte factible.

Finalmente, presentamos a continuación los ítems del formulario que deben completar quienes deseen presentar para su acreditación un proyecto de investigación en el marco del Programa de Incentivos del Ministerio de Educación, Ciencia y Tecnología de la Argentina. Se trata del formato requerido por la Universidad Nacional de La Plata para proyectos a ejecutarse en el período 2007-2009:[8]

1. Identificación del proyecto:
   1.1. Director / codirector.
   1.2. Unidad académica / unidad ejecutora.
   1.3. Denominación del proyecto.
   1.4. Resumen técnico.
   1.5. Palabras clave.
   1.6. Duración del proyecto.
   1.7. Características (tipo de investigación, área, disciplina, campo de aplicación).
   1.8. Transferencia de resultados prevista.
2. Descripción del proyecto:
   2.1. Denominación.
   2.2. Marco teórico o estado actual del tema.
   2.3. Aporte original al tema.
   2.4. Objetivos.
   2.5. Metodología.
   2.6. Metas/resultados esperados en el desarrollo del proyecto.
3. Antecedentes.
4. Aportes potenciales:
   4.1. Contribución al avance del conocimiento científico y/o tecnológico y/o creativo.
   4.2. Contribución a la formación de recursos humanos.
   4.3. Transferencia prevista de los resultados, aplicaciones o conocimientos derivados del proyecto.
5. Plan de trabajo.
6. Recursos humanos intervinientes:
   6.1. Director.
   6.2. Codirector.
   6.3. Investigadores formados.
   6.4. Investigadores en formación.
   6.5. Tesistas, becarios.

[8] Fuente: <www.presi.unlp.edu.ar/secyt/cyt_htm/cyt_file/acred07/plan_acred_07.doc> (fecha de consulta: 23/11/2006).

7. Equipamiento y/o bibliografía:
   7.1. Disponible.
   7.2. Necesario.
   7.3. Fuentes de información disponibles y/o necesarias.
8. Presupuesto del proyecto:
   8.1. Costo mínimo global necesario para llevar a cabo el proyecto.
   8.2. Fondos/recursos disponibles.
   8.3. Fondos/recursos en trámite.
   8.4. Explicitar la factibilidad del plan de trabajo propuesto con los recursos disponibles, en caso de no recibir financiamiento.
9. Aval de la unidad académica.

CAPÍTULO 6

CONCEPTOS DE OBJETO Y DE UNIDAD DE ANÁLISIS. POBLACIÓN Y MUESTRA

6.1. Objeto y unidad de análisis

En el apartado 1.3 se subrayó la importancia que la herramienta intelectual y práctica de la matriz tiene en el "método de la asociación", es decir, en la versión del enfoque estándar que ha dominado en las ciencias sociales.

En estas ciencias se usan varios tipos de matrices, pero el tipo dominante en la fase de la recolección es la llamada "MATRIZ DE DATOS", o matriz "casos por variables". Se decía en el apartado 1.3 que la matriz no es nada más que un cruce entre un haz de vectores paralelos horizontales y un haz de vectores paralelos verticales. En el tipo llamado "matriz de datos", los vectores horizontales se refieren a objetos y los vectores verticales a propiedades de estos objetos.

El término 'OBJETO' se entiende en un sentido gnoseológico, como posible objeto del pensamiento (cualquier cosa en la que se piense). Por lo tanto, los objetos en las filas de una matriz de datos pueden ser individuos (humanos o no), familias, grupos, instituciones, provincias, Estados, eventos, etcétera. Pero, en una matriz dada, todo los objetos en las filas deben ser del mismo tipo: en efecto, no se podría construir una matriz que llevase simultáneamente, por ejemplo, seres humanos y Estados en sus filas, porque los vectores relativos no podrían ser paralelos: es decir, tener referentes del mismo tipo y la misma secuencia de propiedades en las columnas: las propiedades que se pueden referir a individuos no se pueden referir a Estados, y viceversa.

El tipo de objetos que están en las filas determina el tipo de propiedades que pueden estar en las columnas. La propiedad "sexo" no se puede referir a una institución, una provincia o un Estado, como la propiedad "número de ciudadanos adultos" no se puede referir a un individuo.

El tipo de objeto acerca del cual se buscan informaciones en una investigación se llama "UNIDAD DE ANÁLISIS". Esta expresión tiene un referente abstracto, puede ser "ama de casa argentina adulta", pero no puede ser "la señora Ramírez". Las unidades más frecuentemente usadas en la investigación social son el individuo, la familia, el grupo, la empresa, el distrito electoral, el municipio, la provincia, el Estado. También pueden ser unidades eventos, como la elección, la guerra, etcétera.

En una investigación se debe definir no sólo la unidad, sino también el ámbito espacio-temporal que interesa. La necesidad de delimitar el ámbito espacial es obvia: no es difícil darse cuenta de que es cosa diferente estudiar las amas de casa de una ciudad, de una provincia, de un Estado o de un continente. Menos obvia es la necesidad de delimitar el ámbito temporal: como en sociología y ciencia política el sondeo es una herramienta privilegiada de recolección de informaciones, habitualmente se da por sentado que el ámbito temporal sea un genérico presente.[1] Una falta seria de la investigación social es, en efecto, desconocer la perspectiva diacrónica y no aprovechar ni siquiera las oportunidades de analizar los sondeos pasados que se guardan en los archivos de datos.[2]

6.2. Población y muestra

Una vez determinadas la unidad de análisis y el ámbito espacio-temporal, el conjunto de los ejemplares de esa unidad que se encuentran en dicho ámbito es llamado POBLACIÓN.[3] Cada ejemplar de esta población puede devenir un CASO, es decir, el referente de una fila de la matriz.

Cuando la unidad es una provincia y el ámbito un Estado dado en un período dado, o la unidad es un Estado y el ámbito un continente dado en un período dado, la población no es numerosa, y habitualmente se recolectan informaciones acerca de todos sus miembros (es decir, todos los ejemplares de esta unidad dentro del ámbito espacio-temporal). Este procedimiento se llama ENUMERACIÓN COMPLETA.

Pero cuando los miembros de la población son muchos (como en las encuestas sobre individuos adultos de una nación) recoger informaciones sobre todos cuesta un gran esfuerzo e inversión de tiempo y recursos, y sólo se hace raramente, por agencias oficiales del Estado, y para fines que trascienden la investigación social (piénsese en un censo poblacional).

Descontando a esos casos, se presenta el problema de elegir un pequeño subconjunto de estos miembros de la población para investigarlos con un menor gasto de recursos, convirtiéndolos en casos de una matriz de datos. Este

[1] La responsabilidad de esta falta de profundidad diacrónica no cae únicamente sobre el sondeo, porque también los psicólogos —que prefieren usar otras herramientas de recolección además del sondeo— tienen una orientación marcadamente sincrónica.

[2] En los Estados Unidos y Europa del Norte hay muchos archivos de este tipo, donde los investigadores depositan los datos resultantes de sus encuestas para permitir a otros investigadores practicar "análisis secundario" de sus datos. En la Argentina, el INDEC proporciona al público encuestas como la EPH (Encuesta Permanente de Hogares), y muchos estudiantes e investigadores la usan en sus trabajos. Más allá de este ejemplo, sin embargo, en la Argentina —como en muchos otros países— no hay sólidas tradiciones de análisis secundario.

[3] Nótese que en el lenguaje técnico de las ciencias sociales este término tiene un sentido más amplio del que posee en el discurso ordinario. Este sentido fue desarrollado por Malthus (1747) y tiene una gran importancia, no sólo en la estadística sino también en la teoría de la evolución de Darwin.

problema se aborda con una herramienta clásica de las ciencias sociales: el muestreo.

Una MUESTRA es *cualquier* subconjunto, amplísimo o limitadísimo, de miembros de una población que se investiga con el fin de extender a toda la población las conclusiones resultantes del análisis de las informaciones relativas al subconjunto.[4] Esta extrapolación (de los resultados del análisis) de la muestra a la población entera es llamada INFERENCIA ESTADÍSTICA, y tiene reglas precisas que veremos pronto.

Antes cabe resaltar tres malas costumbres muy difundidas en las ciencias sociales actuales. La primera es una costumbre terminológica, y consiste en el hecho de que a menudo se habla de inferencia de la muestra *al universo*. Este uso es impropio, porque el universo es por supuesto infinito mientras que cualquier población sólo puede ser finita. Como pasa a menudo, el uso terminológico impropio no acontece por casualidad, sino porque permite extender a las encuestas de las ciencias sociales fórmulas matemáticas asentadas en supuestos que sólo son legítimos para conjuntos infinitos, es decir universos.

La segunda mala costumbre consiste en el hábito de extender la inferencia más allá de la población de la cual se extrajo la muestra. Un caso clamoroso en la literatura de las ciencias sociales es la renombrada "relación Kinsey". Kinsey era un psicólogo norteamericano que tenía un consultorio privado en una ciudad de Minnesota. Basándose en las declaraciones de sus pacientes (que pueden considerarse una muestra no aleatoria[5] de habitantes de aquella ciudad) él publicó con sus colaboradores dos volúmenes titulados *Conducta sexual del varón humano* (1948) y *Conducta sexual de la hembra humana* (1953). Los sociólogos no están para nada exentos de esta mala costumbre: por ejemplo, el clásico *The People's Choice* (El pueblo elige) está basado en una muestra de 600 electores del condado de Erie en Ohio (Lazarsfeld *et al.* 1944).

6.3. Muestras aleatorias y no aleatorias

La tercera mala costumbre necesita una discusión más profunda. Todas las empresas que producen y venden sondeos —y no sólo ellas— declaran generalmente que sus muestras son "aleatorias y representativas", estos dos términos se usan ritualmente, sin corresponder a ninguna calidad precisa de las muestras mismas ni de los procedimientos que se utilizaron para su extracción. En los pasajes siguientes se dará un significado a estos términos.

Una muestra se dice aleatoria cuando todos los miembros de la población de la que se extrae tienen la misma probabilidad de ser extraídos y entrar en la

[4] En este capítulo se habla del muestreo que se realiza en el marco de investigaciones de tipo estándar (véanse los apartados 2.3 y 2.4). Otras formas de muestreo serán presentadas en el apartado 12.5.

[5] Se aclararán pronto los criterios según los cuales una muestra puede considerarse aleatoria.

muestra. Si se sale a la calle Y de la ciudad X y se entrevista a los primeros cien sujetos que pasan, ésta no es una muestra aleatoria de habitantes de la ciudad X, porque muchos de ellos no pasan jamás por la calle Y, otros pasan raramente y otros diariamente. Por consiguiente, las probabilidades de ser extraídos no son iguales. Para construir una muestra aleatoria de esta población se debe elaborar un catálogo completo de sus miembros y extraer algunos de ellos con una tabla de números aleatorios o con otro procedimiento que garantice efectivamente la misma probabilidad de ser extraído a cada miembro de la población.

Por lo tanto, la naturaleza aleatoria de una muestra depende integralmente del procedimiento de extracción y no tiene nada que ver con su resultado: si se extraen 30 bolillas negras y ninguna blanca de una bolsa, sin mirar de reojo dentro de la bolsa y sin hacer las bolillas distinguibles al tacto, la muestra es perfectamente aleatoria a pesar del resultado, y cualquiera sea la distribución de los colores en las bolillas de la bolsa.

Algunos estadísticos hablan de "muestra aleatoria simple" cuando los miembros de una población tienen la misma probabilidad de ser extraídos, y de muestra aleatoria sin más cuando cada miembro de la población tiene una probabilidad conocida y no nula de ser extraído. Pero esta distinción ofrece una cobertura científica a procedimientos que tienen poco que ver con la ciencia y mucho con los presupuestos de las empresas comerciales de sondeos. Imaginemos una investigación de ámbito nacional. Si la muestra fuese extraída con procedimiento aleatorio, podrían entrar en ella muchos habitantes de áreas remotas de alta montaña o de otros lugares de difícil acceso. Los entrevistadores de las empresas deberían perder tiempo y gastar dinero para contactar a estos sujetos, y para cada una de estas entrevistas la empresa gastaría un múltiplo del valor medio de las entrevistas en las ciudades principales, donde seguramente habitan entrevistadores de su red nacional.

Para reducir esos gastos, las empresas dividen el territorio nacional en áreas, asignando a cada una un número dado de entrevistas. A las áreas remotas se les asigna un número mínimo de entrevistas, de manera tal que todos sus habitantes tengan una probabilidad extremadamente baja, pero conocida y no nula, de entrar en la muestra —y así la ortodoxia estadística es preservada.

Después de ahorrar dinero con esos procedimientos, las empresas tratan de remediar su obvia consecuencia (subrepresentación de todas las áreas periféricas del país) con otro procedimiento discutible: la ponderación. Supongamos que a un área de alta montaña le fueran asignadas 3 entrevistas en lugar de las 27 que le corresponderían según su proporción sobre la población nacional. Debido a que el cociente 27/3 es 9, la ponderación consiste en multiplicar por 9 cada montañés entrevistado, es decir contar sus respuestas 9 veces en todos los análisis que involucren la variable en cuestión.

Se produce así una doble proyección: cada montañés (y cada habitante de áreas remotas) es proyectado en 5, 10, o más de sus clones ficticios. El paso siguiente es proyectar esta mezcla de individuos y de clones a toda la población nacional. En ambos casos se manifiesta un supuesto atomista, en el sentido de que se descuidan no sólo la especificidad de cada individuo, sino también la influencia que el contexto de relaciones sociales en las que cada individuo está inserto ejerce en su personalidad.[6]

Por otro lado, las muestras telefónicas, incluso si han sido extraídas aleatoriamente de una guía de teléfonos, no cumplen ni siquiera con los criterios más tolerantes para ser consideradas aleatorias. Esto se debe al obvio motivo de que todos los miembros de una población que no tienen acceso a un teléfono no tienen ninguna posibilidad de ser incluidos en la muestra. Y por el motivo más sutil de que los miembros de familias numerosas, y/o los que se encuentran raramente en la casa en los horarios en que se acostumbra llamar, tienen *a priori* una probabilidad reducida, y se desconoce el grado de esta reducción para cada categoría.

Hay otros problemas vinculados con el concepto de extracción aleatoria. Un procedimiento que garantice a cada miembro de una población la misma probabilidad de ser extraído es una condición necesaria y suficiente para lograr una muestra aleatoria de dicha población cuando sus miembros son objetos inanimados, que no pueden rechazar ser extraídos y ser examinados. Pero —como todos los entrevistadores saben bien— las poblaciones de seres humanos no cumplen con estos requisitos. A menudo no se encuentran en su casa, ni contestan el teléfono, muchos de los que se encuentran se niegan a la entrevista porque sospechan objetivos comerciales o fiscales, o porque no desean dedicar tiempo a una actividad de la cual no entienden el espíritu ni las finalidades.

Si las probabilidades de no encontrarse en casa o de negarse a la entrevista fuesen igualmente distribuidas en las varias capas de una población humana, estos inconvenientes sólo reducirían las dimensiones de una muestra sin perjudicar su naturaleza aleatoria. Pero casi un siglo de experiencia con los sondeos ha mostrado que los jóvenes y los adultos ocupados tienen mucho menos probabilidades de encontrarse en casa que las amas de casa y los jubilados. Además, adultos empleados y viejos muestran una mayor propensión a negarse a la entrevista. Por lo tanto, estas categorías tienen *a priori* una menor probabilidad de ser efectivamente entrevistadas.

Por consiguiente, aun una muestra perfectamente aleatoria al momento de la extracción se vuelve casi siempre no aleatoria cuando se trata de transformar cada sujeto en un caso de la matriz: es sabido que las amas de casa tienen una mayor probabilidad de ocupar filas en una matriz de datos, seguidas por los adultos desempleados, los jubilados y los jóvenes —en este orden. Los adultos empleados son los que tienen *a priori* una menor probabilidad de convertirse en casos de una matriz de datos.

[6] Por este motivo, muchos investigadores de una escuela muy sensible a la influencia del contexto de relaciones sociales en las elecciones individuales, como el Bureau of Applied Social Research de la Universidad de Columbia, se han mostrado reacios a generalizar más allá del ámbito de sus específicas investigaciones. Véase Martire (2006: cap. 3).

6.4. ¿Representativo de qué?

El otro término fetiche que cabe examinar es 'representativo'. En los textos estándar se leen definiciones como la siguiente: "[una muestra es] representativa si reproduce —en escala reducida— la población objeto del estudio (para permitir la generalización de los resultados obtenidos en la muestra a la población total)" (Corbetta 2003: 159).

Reproducir en escala reducida un diseño o una hoja escrita es algo que una fotocopiadora hace rápida y fácilmente, reduciendo en la misma proporción en la copia las distancias entre cada pareja de puntos en el original. Pero ¿cómo se puede lograr el mismo resultado con poblaciones de seres humanos? Éstas son caracterizadas no sólo por distancias físicas entre sus miembros, sino por muchas otras propiedades. La analogía con la fotocopiadora no puede ser tan simple, pero es útil. En primer lugar, porque resalta el hecho de que la representatividad se juzga confrontando características del original con su análogo en la copia. Esta comparación se puede hacer sólo cuando la copia fue ya producida. Por consiguiente, mientras para juzgar si una muestra es aleatoria debemos mirar el procedimiento con el que se extrae, independientemente del resultado, para juzgar si una muestra es representativa debemos mirar el resultado, independientemente del procedimiento.

Como se decía, para juzgar si la reproducción en escala reducida de un diseño fue correcta, se confrontan parejas de distancias entre puntos. Pero las poblaciones tienen propiedades más importantes que la distancia entre sus miembros. ¿Cómo se confrontan estas propiedades de la población con las de la muestra para juzgar si la última es representativa?

Para responder tenemos que adelantar el concepto de distribución de una propiedad, que va a ser tratado en el capítulo siguiente. Las propiedades de una población pueden ser constantes (como por ejemplo, el sexo en un convento de monjas) o variar. Varían si diferentes miembros de la población tienen diferentes estados en ellas: continuando en el ejemplo, afuera del convento el sexo varía porque algunos individuos son masculinos y otros son femeninos. En este caso, el sexo tiene una distribución, que se puede expresar en cifras absolutas (5.312 mujeres y 4.893 hombres en el pueblo X) o en porcentajes.

Ya tenemos un resultado de la analogía: si se extrae una muestra de los habitantes de ese pueblo, se entiende que ésta no puede ser considerada representativa de la población (es decir, una reproducción de esta última en escala reducida) si los porcentajes de hombres y mujeres son sensiblemente diferentes de los mismos porcentajes en la población.

Pero pronto se plantean dos interrogantes:

a) ¿Cuán diferente es "sensiblemente" diferente?
b) ¿Qué pasa con las otras propiedades distintas del sexo?

Para el primer interrogante los textos de estadística no proporcionan una respuesta. Y eso no es casual: sería un poco ridículo establecer que si hay una diferencia de un punto entre los mismos porcentajes en la población y en la muestra, la última es representativa, y que si la diferencia excede un punto, la muestra no lo es.[7] Esta consideración nos permite sacar una primera conclusión: mientras que para juzgar si una muestra es aleatoria tenemos una pauta clara (misma probabilidad de entrar en la muestra para todos los miembros de la población), para juzgar la representatividad debemos recurrir a consideraciones difusas y subjetivas. Al responder el segundo interrogante veremos que esta difusión y subjetividad caracterizan todo el campo semántico de la representatividad.

Supongamos que la comparación entre las distribuciones de los sexos en la población y en la muestra nos lleve a la conclusión de que son bastante parecidas y, por lo tanto, la muestra es representativa en lo que concierne al sexo, ¿de eso se puede sacar la conclusión de que la muestra también es representativa en lo que concierne a cualquier otra propiedad de la población?

La respuesta es ¡obviamente, no! Y de ésta se siguen algunas importantes consecuencias:

1) La representatividad tiene que ser controlada y eventualmente afirmada para cada propiedad por separado.
2) Debido a que la representatividad se evalúa comparando la distribución de una propiedad en la muestra con la distribución de la misma propiedad en la población, únicamente se puede evaluar para las propiedades cuya distribución en la población es conocida —es decir, sólo para las propiedades que se relevan con un censo poblacional. Para toda otra propiedad, incluyendo opiniones, actitudes, valores, etcétera, la representatividad no se puede mínimamente controlar, y por lo tanto no puede ser afirmada.
3) La oración "esta muestra es representativa" (sin alguna calificación), que a menudo se escucha, no tiene sentido si no se le agrega "en lo que concierne a la(s) propiedad(es) X (Y, Z)". Y además no tiene ninguna credibilidad si no se le agregan tablas que comparan las distribuciones de esas propiedades en la población y en la muestra. Una práctica que —en rigor de verdad— se encuentra muy raramente cumplida en los relatos de sondeos que afirman que la muestra es representativa.

Ahora que se clarificó el significado propio de los dos términos fetiche, se pueden contestar dos preguntas. Empezamos con la más básica: ¿extraer de forma aleatoria una muestra garantiza automáticamente que ésta sea representativa?

La respuesta es, obviamente, negativa, y la hemos adelantado ya, cuando se dijo (véase más arriba en este capítulo) que se pueden extraer de forma perfec-

[7] Cabe resaltar, sin embargo, que en la estadística inferencial hay un sinnúmero de estos umbrales rígidos para discriminar, por ejemplo, si el promedio de la distribución de una variable en una muestra es "significativamente" diferente del promedio de la distribución de la misma variable en la población.

tamente aleatoria (es decir, sin alterar de modo alguno la igual probabilidad de cada bolilla de ser extraída) 30 bolillas negras y ninguna blanca de una bolsa que contiene 30 bolillas negras y 30 blancas (arriba no se especificó el contenido de la bolsa). La muestra es aleatoria, pero ciertamente no es representativa de la población con respecto de la propiedad "color". Paradójicamente, la única manera de garantizar la extracción de un número igual (cualquiera) de bolillas negras y de bolillas blancas sería mirar en la bolsa cuando se extraen —violando abiertamente los requisitos de una extracción aleatoria.

Si uno sale a la calle y entrevista los primeros 50 varones y las primeras 50 mujeres que encuentra, la muestra será *grosso modo* representativa —respecto del sexo— de la población de la ciudad, de la provincia, del Estado, del continente y del mundo. Pero violará abiertamente los requisitos de una extracción aleatoria.

Estos ejemplos pueden naturalmente ser generalizados —con las necesarias adaptaciones— a cualquier otra población, propiedad y forma de extracción. La pregunta consiguiente no puede sino ser: "Si una extracción aleatoria no garantiza representatividad en alguna propiedad, y ésta puede garantizarse en algunas propiedades únicamente con una extracción que dista de ser aleatoria, ¿por qué sería preferible extraer muestras aleatorias?"

La respuesta es que la extracción aleatoria sólo da una garantía negativa, es decir, la de no introducir sesgos de magnitud y dirección desconocida y no controlable en la distribución de las propiedades en la muestra con respecto de las mismas distribuciones en la población. Si salimos a la calle y entrevistamos los primeros 50 varones y las primeras 50 mujeres, la distribución del sexo en esa muestra será *grosso modo* representativa de la población de la provincia, del Estado, etcétera. Pero las distribuciones de muchas otras propiedades relevantes para una investigación social serán ciertamente sesgadas, y a menudo fuertemente sesgadas (piénsese en la distribución del lugar de residencia, de la profesión, de la edad).[8]

Se dijo más arriba que el azar puede producir distribuciones tan sesgadas como 30 bolillas del mismo color de una bolsa con bolillas equidistribuidas en dos (o más) colores. Pero sesgos tan fuertes son extremadamente raros: la mayoría de las extracciones producirán muestras con sesgos pequeños, o incluso ningún sesgo.

Concluyendo, se puede afirmar que la extracción aleatoria no da ninguna garantía absoluta de que la distribución de una propiedad cualquiera en la muestra sea representativa de la distribución de la misma propiedad en la población. Sólo da:

- Una garantía absoluta de que los sesgos (en el sentido de diferencias entre estas dos distribuciones) no sean introducidos por el investigador, sino producidos por el azar.

[8] Serán inevitablemente subrepresentadas todas las ocupaciones que no permiten pasar mucho tiempo en la calle, y los grupos de edad en los que —por una razón u otra— no se pasa mucho tiempo en la calle: los infantes, los viejos, etcétera.

- Una garantía razonable de que esos sesgos sean de magnitud limitada. Si una muestra de alcance nacional es extraída de forma aleatoria, hay una probabilidad prácticamente nula de que todos sus integrantes circulen por el mismo barrio, a diferencia de la muestra que resultaría de salir a la calle a entrevistar a los que pasan con la única preocupación de garantizar la representatividad con respecto al sexo.

6.5. Cómo garantizar una (limitada) representatividad de una muestra aleatoria

¿No hay alguna manera de asegurar algo más que esta garantía negativa con una muestra aleatoria?

Sí la hay, pero con respecto a un número limitado de propiedades, bajo algunas condiciones (que veremos pronto), y sólo con un tipo particular de extracción (denominada SISTEMÁTICA) que vamos a describir.

Supongamos que se tenga un catálogo exhaustivo de los miembros de una población, listado con un orden que no tenga alguna relación con las propiedades que interesan al investigador (por ejemplo, un orden alfabético de apellidos). Supongamos que esta población tiene 80.000 miembros (por ejemplo, los habitantes adultos de una ciudad media) y que queremos extraer una muestra de 400 individuos. Para lograr una muestra aleatoria sistemática se debe:

1) dividir el listado en 400 segmentos, cada cual comprendiendo 200 miembros, número que resulta de dividir el tamaño de la población por el tamaño de la muestra (80.000/400 = 200) y que se denomina "intervalo de muestreo";
2) extraer de forma aleatoria un número de 1 a 200: supongamos que sea el 78.

La muestra sistemática será formada por el 78° individuo de cada segmento, es decir por los individuos que llevan los números 78, 278, 478, y así sucesivamente hasta el número 79.878.

Veamos ahora cómo se puede lograr que esta particular muestra del ejemplo sea no sólo aleatoria sino también representativa de la población de la ciudad con respecto a dos propiedades: el sexo de los habitantes y su barrio de residencia (supongamos que los barrios sean 5, con un promedio de 16.000 habitantes adultos). La condición necesaria y suficiente para conseguir este objetivo es ordenar el listado por barrio, y dentro de cada barrio dividirlo por sexo. Es decir, empezar el listado con todos los varones (adultos) del barrio A, seguidos por todas las mujeres del barrio A, por todos los varones del barrio B, por todas la mujeres del barrio B, etcétera.[9]

Imaginemos que los varones del barrio A sean 9.000. Extrayendo el individuo que lleva el número 78 (o cualquier otro número) en cada segmento de 200

[9] Ese método fue propuesto por vez primera por uno de los autores (Marradi 1997).

individuos, vamos a extraer en la muestra 45 varones del barrio A, es decir el 11,2% de los individuos de la muestra. Esto es exactamente el porcentaje de varones que viven en el barrio A sobre todos los adultos de la ciudad. Se puede fácilmente comprobar que el mismo mecanismo funciona para todos los segmentos en que hemos dividido la población. El resultado es que tenemos una muestra extraída de forma aleatoria que es representativa de la población de la ciudad con respecto al sexo y al barrio de residencia.

Naturalmente, no se puede decir nada de su representatividad con respecto a todas las otras propiedades —es decir, infinitas menos dos. Lo que podemos hacer, si queremos, es segmentar la población teniendo en cuenta también una tercera propiedad (por ejemplo, el nivel de instrucción), e incluso una cuarta (por ejemplo, el grupo de edad). Pero cada vez que se considera una nueva propiedad el número de segmentos a ordenar en secuencia crece en proporción geométrica: si con el sexo y cinco barrios de residencia los segmentos eran 10, considerando cuatro niveles de instrucción ellos serán 40, y considerando también seis grupos de edad, serán 240. Con una población de 80.000 individuos, el tamaño promedio de cada subgrupo (por ejemplo, varones del barrio A con instrucción baja y menores de 25 años) va a ser de 333 individuos. Reduciéndose la diferencia entre el número de individuos que hay en cada segmento (en el ejemplo, 200) y el número de individuos en cada subgrupo, se acrecienta la probabilidad de que la extracción aleatoria sistemática atribuya a un subgrupo un porcentaje de extraídos no exactamente correspondiente a su porcentaje en la población (aun si las diferencias no pueden matemáticamente exceder una unidad en cifras absolutas).

El principal inconveniente de este método es, por lo tanto, la complicación procedimental vinculada con el manejo de subgrupos formados por el producto lógico de categorías en diferentes propiedades. Pero en el caso de que se atribuya gran importancia a la representatividad de la muestra con respecto a pocas propiedades, el método que se ilustró es la manera más simple[10] de conseguirla sin violar la naturaleza aleatoria de la muestra.

[10] Una forma alternativa es el así llamado "muestreo estratificado". Se definen subestratos con un criterio cualquiera y luego se sortean submuestras dentro de cada subestrato. Por otro lado, a menudo se usa un muestreo estratificado "no proporcional" para sobrerrepresentar o subrepresentar un segmento particular de la población estudiada.

Capítulo 7

CONCEPTOS DE PROPIEDADES. VARIABLES, FIDELIDAD Y FIABILIDAD

7.1. La definición operativa

En las columnas de una matriz de datos se encuentran las variables. Una variable es un vector de signos que representan los estados de los casos en las propiedades que interesan. Habitualmente, pero no necesariamente, estos signos son números.[1]

A veces la relación entre un estado en una propiedad y el signo que lo representa es directa e intuitiva: si la entrevistada es Morena Ruiz y tiene 18 años, tal estado en la propiedad "edad" será representado por el mismo 18 en la celda ubicada en el cruce entre el vector-fila relativo a Morena Ruiz y el vector-columna representando la edad. Para saber que este 18 es el número de años de Morena Ruiz, tenemos que saber cuál es el titular del vector-fila y qué propiedad se representa en el vector-columna. Sólo si tenemos estas informaciones, este 18 se convierte de un mero número en un dato —y lo mismo para todos los otros números en la matriz. Pues sólo se puede hablar de una matriz de datos si todos los números (y, más generalmente, los signos) que se ven son interpretados, es decir, son DATOS.

Continuando con Morena Ruiz, es probable que estemos interesados no sólo en su edad, sino también en su título de estudio (en el caso: licenciada). Pero no hay cifras que representen directamente este título (ni los otros títulos).

Por lo tanto, para poner esta información en las celdas relativas necesitamos una convención que conecte el título de estudio "licenciada" a un número particular, y lo mismo para cada otro título que deseemos registrar. Una convención del género se llama PLAN DE CODIFICACIÓN.

Un plan de codificación de la propiedad "título de estudio" podría ser:

[1] Se ponen números para facilitar el análisis estadístico de las relaciones entre variables. Pero este hábito conlleva el riesgo de que se hagan operaciones estadísticas sobre números que sólo son tales en apariencia.

0 = ningún título
1 = sólo educación primaria
2 = sólo educación secundaria
3 = licenciatura
4 = doctorado
9 = información faltante.

Cabe subrayar el hecho de que en cada plan de codificación se necesita prever un número particular para informaciones faltantes. Para esta función de señalar la falta de informaciones se elige habitualmente un número al margen del abanico de los otros: un 0 o un 9 si se usan números de una cifra, un 0 o un 99 si se usan números de dos cifras, y así sucesivamente.

Se puede comentar que el plan de codificación que hemos visto es bastante grosero —en el lenguaje metodológico, poco sensible. La SENSIBILIDAD es un cociente entre el número de estados en una propiedad que consideramos como categorías en nuestro plan de codificación y el número de estados diferentes que tiene la misma propiedad.[2]

Para acrecentar la sensibilidad podemos considerar una gama más amplia de estados en nuestro plan de codificación; por ejemplo:

0 = ningún título
1 = sólo educación primaria
2 = sólo educación secundaria
3 = sólo educación terciaria
4 = licenciatura
5 = especialización
6 = maestría
7 = doctorado
8 = posdoctorado
9 = información faltante.

El plan de codificación es parte de un conjunto de reglas y convenciones[3] que permiten convertir una propiedad de los objetos o sujetos en la realidad en una variable en la matriz de datos. Este conjunto se llama DEFINICIÓN OPERATIVA y las partes que lo integran varían según la naturaleza de la propiedad. La variación más considerable está vinculada al tipo de unidad de análisis, y por consiguiente al tipo de técnica de recolección.

[2] Cuando desconocemos el número de estados diferentes, obviamente el cociente queda indeterminado. Pero el concepto de sensibilidad mantiene su utilidad: sea cual sea este número, agregando una categoría se acrecienta la sensibilidad, y eliminando una categoría la sensibilidad disminuye.

[3] Se entiende que estas reglas y convenciones son específicas no sólo de cada propiedad, sino también de cada investigación, en el sentido de que dependen de la elección del o de los investigadores. Como observa Corbetta (2003: 90), en las ciencias físicas "se forman dentro de la comunidad científica convergencias que permiten llegar a definiciones de carácter intersubjetivo". Pero en las ciencias sociales estamos muy alejados de esta situación.

Si queremos recoger informaciones acerca del título de estudio de los entrevistados, la definición operativa incluye:

1) el texto de la pregunta relativa (ejemplo: "¿Cuál es su título de estudio?");
2) un plan de codificación como el que hemos visto arriba;
3) el hecho de que este plan sea directamente planteado al entrevistado junto con la pregunta pidiéndole que elija una de las respuestas alternativas que se le ofrecen, o sólo sea empleado como una lista de voces a las cuales atribuir las respuestas espontáneas de cada entrevistado a la pregunta relativa;
4) la indicación del vector-columna en que se deben poner las informaciones acerca del título de cada entrevistado —informaciones que se convirtieron en datos representando cada una de ellas con el correspondiente número del plan de codificación;
5) si el investigador sospecha que un número bastante alto de entrevistados puede mentir acerca de su título, y si él dispone de los recursos necesarios (tiempo, fondos, colaboradores confiables), puede también insertar en la definición operativa un procedimiento con el que se controlará la sinceridad de las declaraciones de los entrevistados —por ejemplo, recurriendo a documentos oficiales.[4]

Veamos ahora como se presenta la definición operativa cuando la unidad de análisis es una provincia u otro agregado territorial: imaginemos que la propiedad que interesa sea la proporción de población licenciada en cada provincia. En este caso —como para la mayoría de las propiedades relativas a provincias, Estados u otros agregados territoriales— se recurre al censo.

La definición operativa indicará:

1) las tablas de la publicación del censo donde se encuentran informaciones sobre el número de licenciados y el número de habitantes con más de 20 años de edad;
2) la operación matemática que se debe realizar (una división del número de licenciados por el número de habitantes de la edad requerida);
3) la ulterior operación matemática que se realizará para hacer más legible el resultado (por ejemplo, multiplicar los cocientes por 1.000 y redondear el resultado para eliminar los decimales);
4) la indicación del vector-columna en que se debe poner el número resultante;
5) un procedimiento de control mecánico, por ejemplo, la ejecución de los cálculos al revés para ver si se obtienen los mismos números que aparecen en las tablas de las que se iniciaron los cálculos;
6) una valoración de la plausibilidad de los resultados. En efecto, puede pasar que la proporción de población licenciada en una provincia rural

[4] Acerca del problema del control de la fidelidad de los datos, véase el apartado 7.3.

y marginal aparezca más alta que la misma proporción en una provincia urbanizada y terciarizada. No se puede excluir que ésta sea la situación efectiva; pero su escasa plausibilidad sugiere revisar los cálculos e incluso poner en duda los mismos datos en las tablas del censo.

En los ejemplos hemos presentado las partes de dos definiciones operativas que eran específicas de las propiedades en cuestión (título de estudio individual y proporción de población licenciada en la provincia). Pero si consideramos que una definición operativa es el conjunto de todas las operaciones, intelectuales y materiales, que permitan pasar de estados reales a números (en un vector-columna) que representen esos estados en la matriz, debemos concluir que, junto con las reglas y convenciones específicas de la propiedad en cuestión, una definición operativa incluye reglas y convenciones más generales y comunes a todas las propiedades que se indagan en esa investigación específica y en todas aquellas análogas. Por ejemplo, si la técnica de recolección de los datos es una entrevista con cuestionario, la definición operativa de cada propiedad investigada incluye el arte de redactar el texto de las preguntas de manera simple, el arte de ordenarlas en el cuestionario de forma correcta y eficaz, el arte de contactar al entrevistado y plantearle preguntas obteniendo respuestas sinceras, y, por lo tanto, el nivel de adiestramiento y motivación de los entrevistadores —en pocas palabras, toda la sabiduría que se ha acumulado en décadas de desarrollo de los sondeos.

La tesis de que estos aspectos propios de una técnica de recolección en general pertenecen a la definición operativa de cada propiedad investigada con esa técnica puede parecer extraña si no se reflexiona sobre el hecho de que una pregunta mal formulada o mal planteada puede obtener respuestas radicalmente diferentes de las que obtendría si fuera formulada de la mejor manera, y que la misma pregunta sobre una cuestión embarazosa obtiene a menudo respuestas diferentes según se ponga al inicio, a medio camino o al final de un cuestionario.[5]

7.2. La fidelidad de los datos

Un dato tiene escasa fidelidad si no representa correctamente el correspondiente estado en la propiedad *según las convenciones de la definición operativa*.[6] Como hemos visto antes, una causa de datos de escasa fidelidad son las declaraciones mentirosas de los entrevistados: se tiende a reducir la

[5] Décadas de investigaciones específicas han averiguado de forma conclusiva los efectos determinantes de variaciones en el texto de una pregunta, en su posición en un cuestionario, en el número y sucesión de las alternativas de respuesta, en el modo en que se presentan, etcétera.

[6] Para aclarar lo que significa la oración en cursiva, piénsese en una definición operativa que prescribe registrar la propiedad edad en décadas; en este caso el número que representa una edad de 10 años y una de 19 será el mismo (supongamos, el número 1). En eso, se tratará de una definición operativa con un plan de codificación poco sensible; pero los datos 1 que corresponden a edades entre 10 y 19 años representarán la edad en la forma prevista por la definición operativa.

edad y el ingreso, a elevar el título de estudio y la frecuencia con la que se leen libros, se va al teatro y se siguen programas culturales por televisión. Raramente se admiten hábitos sexuales heterodoxos, simpatías por partidos y grupos extremos, conductas ilegales (evasión impositiva, ausentismo en el trabajo, etcétera).

Los investigadores que estudian habitualmente las opiniones de la gente saben por experiencia que los entrevistados —si no tienen un interés particular y opiniones firmes en el tema de la pregunta— tienden a declarar como propias las opiniones que creen ortodoxas y compartidas por la mayoría. Este fenómeno, que no se puede considerar una mentira en sentido estricto, es conocido como "deseabilidad social", y es tan difundido que hace poco fiables los resultados de encuestas sobre opiniones realizadas con preguntas sencillas y directas, ya que es fácil para el entrevistado imaginar cuál sea la respuesta convencional y por lo tanto "socialmente deseable".[7]

Además, los entrevistados pueden dar involuntariamente informaciones que generan datos no fidedignos. Puede ser que no entiendan la pregunta o no le atribuyan el mismo significado que le da el investigador. Puede ser que la entiendan correctamente pero se equivoquen acerca de su propio estado en la propiedad investigada. Puede ser que entiendan mal el mecanismo de algunas técnicas sofisticadas que imaginaron psicólogos y metodólogos para registrar opiniones y actitudes[8] y alteren sus estados sin quererlo.

Si la pregunta implica el recuerdo de acciones y situaciones pasadas, la experiencia ha mostrado que la memoria presenta algunos sesgos sistemáticos, en el sentido de que se tiende a acercar en el tiempo algunos tipos de eventos y a alejar otros (véase Montesperelli 2005: cap. 2).

Naturalmente, el entrevistado no es la única fuente de datos no fidedignos en un sondeo. El entrevistador, a su vez, puede leer mal algunas preguntas o algunas respuestas del plan de codificación; puede influir al entrevistado manifestando sus propias preferencias o sus expectativas acerca de las respuestas; puede equivocarse asignando una categoría errada a una respuesta del entrevistado.

Faltas materiales de trascripción se pueden producir en cada eslabón de la cadena de pasajes que une la respuesta vocal del entrevistado a una cifra en el soporte electrónico que representa la matriz de datos.

Se podría pensar que estas faltas materiales de trascripción son la única causa posible de escasa fidelidad cuando las informaciones se refieren a unidades de análisis como provincias o Estados, y son recogidas por censos o encuestas de departamentos estadísticos públicos. Pero no es así: la gran mayoría de estos datos que convencionalmente se llaman "duros" —para distin-

[7] El problema de la "deseabilidad social" fue identificado por primera vez por Allen Edwards (1957b) analizando las respuestas al afamado test Minnesota Multiphasic Personality Inventory. Otros importantes estudios del problema son Wiggins (1968), Orvik (1972), DeMaio (1984), Presser y Traugott (1992), Corbetta (2003).

[8] Véase el capítulo 8.

guirlos de los datos de sondeo, que serían "blandos"— provienen de entrevistas, y por lo tanto sufren de las mismas causas de infidelidad que hemos examinado antes.[9]

7.3. Para controlar la fidelidad

Dado que la escasa fidelidad de los datos empeora la calidad y reduce la credibilidad de una investigación, se debería dedicar mucha más atención al control de la fidelidad de la que se dedica habitualmente.[10]

En el apartado 7.1, hablando de la definición operativa, ya se hizo referencia a algunas formas de ese control. Ahora se reseñará la gran variedad de estas formas.

La más rápida —que no cuesta nada de dinero e insume poco tiempo, y por lo tanto no hay alguna razón para evitar— se desarrolla explotando la matriz misma,[11] y puede seguir dos caminos. El más simple es una inspección de las distribuciones de frecuencia de cada variable, es decir, de cuántos datos tiene cada uno de los códigos numéricos posibles en esa variable.[12] De todas formas, esta inspección es el primer paso del análisis, porque nos permite tomar contacto con los datos que hemos recogido. Además, nos permite descubrir si algunos datos tienen códigos que no estaban previstos en el plan de codificación (por ejemplo, un dato con código 3 en la variable "sexo" cuando el plan prevea asignar 1 a las mujeres, 2 a los varones y 9 a las informaciones faltantes).

El análisis de las distribuciones de frecuencia de cada variable de una matriz para encontrar códigos no previstos se llama *wild code check* (control de códigos "locos"). Si la unidad de análisis es una provincia y la variable en cuestión es el número de habitantes, o de licenciados, o de hospitales, ese control se realiza de forma diferente, buscando a qué provincias corresponden los valores más altos, o más bajos, en cada variable cuantitativa, para ver si las cifras que encontramos condicen más o menos con lo que sabemos de esa provincia: por ejemplo, si resulta que La Rioja tiene más licenciados que la provincia de Buenos Aires, esto es una señal de que hubo una falta material en la codificación de los datos (por ejemplo, se agregó un cero al final de la cifra, con el efecto indeseado de multiplicarla por diez).

El otro camino con el que podemos controlar la fidelidad de (algunos) datos

[9] Muchos de los que han trabajado con estos datos "duros" denuncian que son escasamente confiables. Véanse por ejemplo Taylor y Hudson (1972) y los ensayos recogidos en Merritt y Rokkan (1966).

[10] En efecto, como se observa en un texto clásico de epistemología, "no hay manipulación matemática que pueda remediar una escasa precisión de los datos que se manipulan" (Cohen y Nagel 1934: 292).

[11] Sin embargo, si bien estos controles rápidos en la matriz permiten detectar datos no fidedignos, no permiten descubrir cuáles deberían ser los datos fidedignos. Para lograr esto, tenemos que reexaminar el proceso de recolección, e incluso repetirlo.

[12] Acerca de las distribuciones de frecuencia, véase el apartado 8.1.

explotando la matriz misma se basa en el hecho de que algunas combinaciones de categorías son social o jurídicamente imposibles en una sociedad dada: no puede haber curas católicos de sexo femenino, ni jueces que no tengan el título de abogado, ni electores menores de dieciocho años, etcétera.

Como se puede entender, esta forma de control, que se llama *consistency check* (control de coherencia), sólo se puede aplicar a un número limitado de situaciones —lo que de todos modos no es una buena razón para no aplicarlo.

Todas las otras formas de control se aplican confrontando los datos con informaciones que se encuentran fuera de la matriz. Ya hemos visto un caso hablando en el apartado 7.1 de la valoración de la plausibilidad de los resultados de una definición operativa cuando la unidad de análisis es una provincia u otro agregado territorial. En el mismo apartado se vio que el recurso a documentos oficiales permite controlar algunas declaraciones de los entrevistados (por ejemplo, declaraciones sobre el hecho de que votaron en las últimas elecciones, o que tienen o no tienen teléfono, licencia de conducir, etcétera).[13]

Como el precedente, ese tipo de control es conceptualmente simple y tiene un resultado en principio fiable. Pero necesita una labor paciente y tediosa y una considerable inversión de tiempo y dinero, porque los documentos oficiales no siempre son tan fáciles de consultar como se proclama oficialmente. Además, con la difusión del culto de la privacidad en los países occidentales, el número de documentos oficiales que se pueden consultar se ha ido reduciendo dramáticamente. De todas formas, este tipo de control sólo puede aplicarse a estados en propiedades que sean oficialmente registradas —seguramente una minoría entre las que interesan al investigador social.

Un tercer tipo de control no sufre de esta limitación, en el sentido de que puede aplicarse a cualquier propiedad individual. Pero necesita grandes inversiones de tiempo y dinero y —a diferencia de las otras formas— una gran destreza y sabiduría de parte de los que ejercen los controles.

Según nuestro conocimiento, esta forma de control —imaginada, por ejemplo, por McNemar (1946)— fue realizada sistemáticamente sólo en un centro de investigaciones metodológicas de la Universidad de Lodz, en Polonia, liderado por Jan Lutinski, que la llamó "entrevista sobre la entrevista". La situación en la que se desarrolló esta técnica de control es muy particular y merece ser analizada en detalle.

Es sabido que los regímenes autoritarios o totalitarios son hostiles a las ciencias sociales, porque tienden a dar una conciencia crítica a la población dificultando su adoctrinamiento. Además, como el marxismo sostenía ser la única ciencia legítima de la sociedad, después de la toma del poder en Rusia, los comunistas procuraron erradicar a las nacientes ciencias sociales.[14]

[13] En los Estados Unidos y en algunos otros países se puede controlar el ingreso declarado por un entrevistado con su declaración fiscal (Katona 1951).

[14] Suerte radicalmente diferente tocó a la psicología, que fue impulsada para crear el nuevo "hombre soviético". Con Vygotskij, Lurija, Leontiev y otros, la psicología rusa estuvo en efecto a la vanguardia en el período de entreguerras.

Lo mismo hicieron al final de la Segunda Guerra Mundial en Europa oriental; con la excepción de Polonia, donde la sociología tenía gran tradición y alto prestigio.[15] Los gerentes comunistas polacos no se arriesgaron a callar y desterrar a los sociólogos y en cambio trataron de explotar sus conocimientos y su sabiduría para conocer mejor lo que la población pensaba realmente del régimen.

En aquel tiempo se decía comúnmente que cada polaco tenía tres verdades: una para el partido, una para la Iglesia y una para sus amigos de confianza. Obviamente, los sondeos que el partido organizaba periódicamente para conocer el estado de la opinión pública sobre tal o cual cuestión sólo podían brindarle respuestas que reproducían exactamente las verdades oficiales, y por lo tanto tenían un valor informativo nulo. Los políticos se daban cuenta de que para tener un cuadro más realista debían permitir a los sociólogos investigar lo que realmente pensaba la gente.

En este contexto se desarrolla la actividad del Centro Metodológico de Lodz (la tercera ciudad de Polonia), una institución pública cuya única tarea era refinar las herramientas del sondeo para mejorar la fidelidad de los datos. Además, el centro no sufría ninguno de los condicionamientos que caracterizan la actividad del metodólogo (a menudo presupuestos insuficientes, tiempo siempre limitado, presiones para publicar rápidamente, tal vez sutiles presiones en dirección a los resultados deseados por los que encargan o por el teórico).

Después de organizar un sondeo sobre un problema determinado, el centro seleccionaba un pequeño subconjunto de los entrevistados y encomendaba cada uno de ellos a uno de sus investigadores, que debía dedicarle todo el tiempo necesario para ganar su confianza, invitándolo entre otras cosas a comentar libremente los temas del cuestionario para llevarlo a dar respuestas más sinceras que las dadas al cuestionario estandarizado (Pilichowski y Rostocki 1978).[16] Lo más difícil era obviamente convencer a los entrevistados de que no corrían ningún riesgo como consecuencia de lo que decían. Esto estaba garantizado por los líderes políticos, que no tenían interés en cortar esos valiosos canales de informaciones confiables, y por lo tanto respetaban el estricto anonimato que los investigadores a su vez garantizaban a los entrevistados.

Hemos dedicado gran atención a la actividad de este centro polaco para mostrar qué particulares y extrañas condiciones deben darse para posibilitar un trabajo en profundidad acerca de la fidelidad de los datos de encuestas.[17]

[15] Piénsese en Ossowski, Chalasinski, Znaniecki.

[16] El metodólogo italiano Gilli propuso efectuar esta profundización durante la misma entrevista inicial (1971: 117). Pero de esta forma ésa se convertiría en una entrevista en profundidad —una herramienta típica de la aproximación no estándar— (véase el capítulo 12).

[17] En verdad, algunos controles de fidelidad de las respuestas estandarizadas a través de sucesivas preguntas abiertas se realizan en algunos centros de investigaciones en Inglaterra y Estados Unidos (véanse por ejemplo Schuman 1966; Belson 1981). Son controles llevados a cabo a menudo durante la misma entrevista y siempre por entrevistadores —no por investigadores como en Polonia. Nada que ver con el empeño y la profundidad del trabajo del Centro de Lodz.

No es nuestra intención proponer este camino para controlar la fidelidad de centenares de respuestas a los millares de cuestionarios estandarizados de un sondeo de masas. El tiempo, el dinero y la habilidad requeridos para plantear preguntas en una entrevista estandarizada son inconmensurables con la inversión de tiempo, dinero y habilidad necesaria para controlar la fidelidad de las respuestas en el modo que hemos descrito.

Tampoco se quiere afirmar que una "entrevista sobre la entrevista" a la manera del Centro de Lodz va a garantizar la fidelidad de las respuestas a preguntas sobre opiniones y actitudes. Incluso los investigadores más hábiles y persuasivos no pueden garantizar esto.

Lo que se pretende afirmar es que éste es el único camino para controlar la fidelidad de respuestas a preguntas sobre opiniones y actitudes, cuando por fidelidad se entienda la correspondencia entre el dato que se anota en la matriz y lo que el entrevistado opina realmente. Las herramientas de control de la fiabilidad que los psicólogos han desarrollado durante el siglo pasado, y que se usan en todas las ciencias sociales, no logran este resultado, a pesar de sus pretensiones. Trataremos de mostrarlo en el apartado siguiente.

7.4. Los coeficientes de fiabilidad: una crítica

El largo camino que conduce a la familia actual de coeficientes que tratan de medir la fiabilidad (en inglés *reliability*) empieza en la astronomía del siglo XVIII.[18] En dicho siglo la consolidación del paradigma kepleriano-galileano-newtoniano, con su robusto *corpus* teórico, permitió calcular en abstracto y con gran seguridad la posición de los planetas conocidos entonces. Confrontando estas posiciones teóricas con las posiciones observadas mediante telescopios, los astrónomos se dieron cuenta de que hay siempre diferencias entre unas y otras. Dado que tenían plena confianza en sus teorías, no dudaron de que esas diferencias se debieran a fallas de las herramientas y/o de los hombres que las manejaban.

Nace así el problema de la precisión de las herramientas y de los observadores —que se convertirá en un problema central en la filosofía de la ciencia. El primer ensayo sobre "fallas de observación" es un artículo de 1757 del matemático inglés Thomas Simpson. Poco después, al problema se dedicaron dos grandes matemáticos y astrónomos franceses, el conde de Lagrange y el marqués de Laplace. A principios del siglo XIX el matemático alemán Karl Friedrich Gauss recolectó unas series de observaciones de la posición de los planetas registradas durante años por varios observatorios astronómicos europeos. Confrontando las posiciones de un planeta cualquiera observadas por varios telesco-

[18] En realidad, parece que los astrónomos sumerios usaban el promedio de las posiciones observadas de una estrella como estimación de su posición efectiva. Eso muestra la gran sofisticación gnoseológica de esos sabios, plenamente conscientes de la posibilidad de un desajuste entre la realidad y lo que se observa de ella.

pios con su posición teórica, calculada según le teoría astronómica, Gauss encontró que las posiciones observadas se disponían alrededor de la posición calculada, y que había muchas diferencias pequeñas y pocas diferencias grandes en ambos lados de dicha posición. En otras palabras, las fallas de observación tenían una distribución campanular, que tomará el nombre de ese matemático (distribución gaussiana).[19]

Si se suman los cuadrados[20] de estas diferencias entre las posiciones observadas y la posición calculada, se puede cuantificar el nivel de (im)precisión del complejo de las herramientas y operadores. La medida de esta imprecisión es exactamente la varianza de las observaciones, es decir, la media de sus desviaciones cuadradas de la posición calculada.[21] Obviamente, el inverso de esta varianza mide la precisión.

Nótese que ésta es una medida global, que se refiere —como se decía— al complejo de herramientas y observadores. Para cuantificar la precisión de cualquier pareja telescopio/observador se debería sólo considerar sus observaciones, y para atribuir la falla a uno u otro se deberían alternar varios observadores en el mismo telescopio y varios telescopios con el mismo observador. Pero los astrónomos no han seguido este camino.

En el siglo XIX, la idea de confrontar los resultados de observaciones repetidas del mismo fenómeno para cuantificar la precisión de las herramientas de medición (por ejemplo, la balanza) se difunde en las ciencias físicas. Al final del siglo, el psicólogo alemán Hermann Ebbinghaus trata de aplicar la idea en sus investigaciones sobre la memoria (1885). Poco después, el gran psicólogo inglés Charles Spearman, en el mismo ensayo (1904) en que esboza los fundamentos del análisis factorial, propone medir la fiabilidad de un test calculando el coeficiente de correlación[22] entre dos vectores de puntajes obtenidos en una muestra de sujetos a los que el test fue aplicado en dos momentos diferentes. Se trata de la primera codificación de la técnica conocida como control de fiabilidad *test-retest*.

Puesto que esta técnica es todavía regularmente usada en psicología, se describirá en detalle su aplicación:

a) para "medir" la fiabilidad de un test (que comprende habitualmente un cierto número de distintas preguntas o pruebas) se le aplica en un cierto día a una muestra de sujetos (habitualmente, los estudiantes de la clase universitaria de psicología dictada por el psicólogo que propone el test);[23]
b) se registran las respuestas a las preguntas (o las soluciones de las pruebas) y se le atribuye un puntaje a cada respuesta;
c) a cada sujeto se le atribuye la suma o el promedio de sus puntajes, y esa cifra, siendo el resultado global de cada individuo en el test, se pone en un vector-columna (un puntaje para cada sujeto);
d) después de algún tiempo (habitualmente, una o más semanas) se repite la fase (a) con los mismos sujetos, las mismas preguntas o pruebas y en una situación en cuanto posible análoga;
e) se repite la atribución de puntajes a los sujetos, como sub (b) y (c), y se ponen las cifras en otro vector-columna;
f) se calcula un coeficiente de correlación entre los dos vectores;
g) esta cifra, llamada "coeficiente de correlación test-retest", se considera una medida de la fiabilidad del test en consideración.

Muchos psicólogos y psicómetras han expresado varias críticas a esta forma de control de la fiabilidad:

a) en el intervalo de tiempo entre el test y el retest, el estado de algunos sujetos en la propiedad investigada puede efectivamente cambiar, como pueden cambiar algunos aspectos de la conducta del investigador que aplica el test, o de la situación en general.[24] Todo eso va a reducir el coeficiente de correlación y por consiguiente nuestra estimación de la fiabilidad del test; para limitar este sesgo sería por lo tanto aconsejable un intervalo muy corto entre test y retest;
b) por otro lado, la primera suministración puede cambiar artificialmente el estado del sujeto en la propiedad, y esto en varias formas:
   1) el sujeto puede acordarse las respuestas que dio en la primera ocasión y repetirlas en la segunda para mostrarse coherente:[25] esto va a acrecentar la aparente fiabilidad del test, y para limitar el sesgo sería aconsejable un intervalo muy largo entre test y retest;
   2) si el test concierne a cualquier forma de habilidad, la primera suministración puede actuar como entrenamiento, y la prestación de algunos sujetos puede mejorar en el retest:[26] esto va a reducir la aparente fiabilidad del test, y para limitar el sesgo sería aconsejable un intervalo largo entre test y retest;
   3) la primera suministración del test puede acrecentar el interés de algunos sujetos hacia los problemas investigados, y ellos pueden re-

[19] La particular forma de distribución campanular que se llama distribución normal había sido definida en abstracto por el matemático francés Abraham De Moivre (1733). Sin embargo, Gauss fue el primero en encontrar esta importante aplicación práctica.

[20] Gauss introduce el criterio de sumar los cuadrados de las desviaciones porque, si no se eliminase el signo con la elevación al cuadrado, las desviaciones negativas tenderían a anular las desviaciones positivas. Desde entonces este criterio ha sido la regla en estadística.

[21] Sobre el concepto de varianza, véase el apartado 8.3.

[22] Esta fundamental herramienta del análisis estadístico (véase el apartado 15.6) había sido recientemente propuesta por Francis Galton (1888) y perfeccionada por Karl Pearson (1896). Sobre este período de la psicología británica, fundamental para el desarrollo de las actuales técnicas de análisis de datos en el enfoque estándar, véase Piovani (2006).

[23] La misma observación en Corbetta (2003: 153).

[24] Véanse, por ejemplo, Webb *et al.* (1966), Converse (1970), Berka (1983).

[25] Véanse, por ejemplo, Kuder y Richardson (1937), Ingram (1977).

[26] Véanse, por ejemplo, Cronbach (1949), Anastasi (1953), Corbetta (2003).

flexionar, informarse, etcétera[27] —lo que puede cambiar sus respuestas a las preguntas y por consiguiente reducir la aparente fiabilidad del test;

4) por lo demás, la primera suministración del test puede tener un abanico de otros efectos[28] (acostumbramiento y desencanto, menor miedo y por lo tanto mayor sinceridad, mayor control de la situación y por lo tanto mejor capacidad de adivinar las respuestas "socialmente deseables", etcétera).

Por este complejo de motivos Cronbach (1947), Anastasi (1954) y otros aconsejan señalar siempre el intervalo entre la primera y la segunda suministración del test —lo que sería determinante si los estados sólo cambiasen espontáneamente, como función lineal del tiempo. Pero hemos visto que no es así: hay muchos otros factores de distorsión, que varían de situación a situación y que no se pueden determinar fácilmente —y todavía menos controlar.

Como se vio, el problema medular que —de una forma u otra— llamaba la atención crítica de algunos psicólogos era la distancia temporal entre las dos suministraciones del test. En las primeras décadas del siglo XX, los esfuerzos de los psicómetras apuntaban a eliminar esta distancia, y con ella todos los problemas que traía aparejados.[29] ¿Pero cómo lograr igualmente dos vectores de cifras para correlacionarlos y calcular un coeficiente, con su aspecto tan sólido y científico?

Una solución hallada fue realizar dos tests sobre el mismo tema, con el mismo número de preguntas o pruebas, formuladas de manera análoga, que se bautizaban como "paralelos". Los dos vectores deseados se producían obviamente sumando los puntajes logrados por cada sujeto al contestar las preguntas (o solucionando las pruebas) de cada test. Si la correlación resultaba alta, eso confirmaba el supuesto de equivalencia de los dos tests y simultáneamente la fiabilidad de cada uno y del conjunto de los dos.[30]

A principios de la década de 1930 los psicómetras propusieron una técnica aún más sencilla: administrar un solo test y luego crear los dos vectores deseados simplemente sumando por separado los puntajes conseguidos por un sujeto en las preguntas que tienen un número par en la secuencia del test y en las que tienen un número impar.[31] El coeficiente de correlación así producido fue bautizado *split-half* (subdivisión a mitad) *reliability coefficient*.

Algunos críticos[32] han notado que la división entre preguntas con número par y preguntas con número impar es arbitraria, y que cambiando el criterio de

[27] Véanse, por ejemplo, Campbell y Stanley (1963), Bohrnstedt (1970), Carmines y Zeller (1979).

[28] Véase, por ejemplo, Scott (1968).

[29] Incluso el problema práctico de convencer a los sujetos testeados la primera vez de dejarse testear una segunda vez con una herramienta que habitualmente es larga y aburrida, por motivos que veremos pronto.

[30] Sobre esta técnica, bautizada *parallel forms*, véanse Thurstone (1928), Anastasi (1954).

[31] Véanse Brownell (1933), Zubin (1934), Kuder y Richardson (1937).

[32] Kuder y Richardson (1937), McKennell (1977), Giampaglia (1985).

división interna el coeficiente *split half* podría cambiar de modo no desatendible. Pero la mayoría de los psicómetras tienen una preocupación diferente: dividiendo por dos el número de preguntas (o pruebas) se reduce automáticamente el nivel del coeficiente.[33] Se adopta con entusiasmo una fórmula que permite "corregir" (es decir: aumentar) el valor del coeficiente neutralizando el efecto de la reducción del número de preguntas/pruebas.

Quien tiene experiencia acerca de las reacciones de los sujetos investigados a frases planteadas en forma de batería sabe que muchos tienden a acordar con ellas sin reflexionar, independientemente de su sentido, por motivos que van de la prisa a un malentendido respecto de la institución que encarga el estudio, al deseo de quitarse un peso de encima. Este fenómeno —conocido como *response set*— es un mal endémico de las baterías, y acerca de sus causas existe una literatura vastísima.[34] Obviamente no hay ninguna garantía de que una respuesta que un sujeto da para quitarse un peso de encima corresponda a sus opiniones efectivas. Pero cada sujeto que acuerda con todas las oraciones de una batería va a aumentar inevitablemente el nivel de cualquier coeficiente asentado en la correlación entre vectores de cifras.[35]

Este patente absurdo no parece preocupar a ninguno: si el único problema es acrecentar el nivel de un coeficiente de correlación entre dos vectores de cifras en la matriz de datos,[36] todo lo que coadyuve a este objetivo es bienvenido. Nadie se plantea el problema de la correspondencia de dichas cifras con los estados efectivos de los sujetos investigados.[37] El mundo real ha desaparecido. Un metodólogo sueco lo ha declarado simple y abiertamente: "La acumulación de estos estudios nos ha enseñado a confiar en estas técnicas [...]. Usándolas, los investigadores ya no tienen que encarar el largo y aburrido proceso de controlar la fiabilidad de sus datos" (Zetterberg 1954: 124).

Un progreso meramente técnico se produjo a fines de la década de 1930, cuando dos integrantes del más importante laboratorio psicométrico nortea-

[33] Muchas fórmulas de coeficientes que "miden" la fiabilidad de un test (por ejemplo el alfa de Cronbach) son construidas de tal forma que el nivel del coeficiente depende también del número de preguntas o pruebas: por este motivo muchos tests psicométricos son largos y aburridos. Para aumentar artificialmente el nivel del coeficiente se somete a los estudiantes "cautivos" a una faena que a muchos puede parecer sin sentido. Lo han observado, entre otros, McKennell (1970), Davies (1977), Ferrand y Martel (1986).

[34] Mencionamos por ejemplo Cronbach (1946), Block (1965), Hui y Triandis (1985), Greenleaf (1992), Johnson *et al.* (2005), y los ensayos en Berg (1967) y en Dijkstra y Van der Zouwen (1982).

[35] "Un alto valor del coeficiente puede depender totalmente de factores de sesgo" (Bartlett *et al.* 1960: 703).

[36] La idea de producir dos vectores para correlacionarlos y obtener un coeficiente se difunde de la psicología a otras disciplinas. Como en el *test-retest* se variaba el tiempo y en el *split half* y "formas equivalentes" se variaba la herramienta, se pueden variar los observadores que valoran un texto cualquiera (por ejemplo, en el análisis de contenido) y obtener un coeficiente de acuerdo entre codificadores (*inter-individual* o *inter-coder reliability coefficient*: Zetterberg 1954; Singer 1982).

[37] Por lo menos algunos psicólogos admiten que en estos casos no se puede hablar de fiabilidad, y en cambio bautizan "coeficiente de estabilidad" al *test-retest* y "coeficiente de coherencia interna" al *split-half* (Cronbach 1947; Anastasi 1954; McKennell 1970).

mericano (Kuder y Richardson 1937) propusieron una técnica que superaba el límite que compartían todos los coeficientes formulados hasta entonces, es decir basarse en una correlación entre dos vectores solos. Con este nuevo coeficiente se calculaba la correlación entre todas las respuestas a las preguntas o pruebas de un test, con tal de que ellas fuesen dicotómicas. Esta limitación fue superada en la segunda posguerra, cuando Cronbach (1951) propuso el coeficiente *alpha*, aplicable a todo tipo de batería. En las décadas siguientes la evolución técnica siguió siendo rápida: entre otros, se han formulado coeficientes basados en el análisis factorial (Heise y Bohrnstedt 1970) y en el análisis de componentes principales (Armor 1974).[38]

Un punto a destacar es que desde su inicial concepción por Spearman, esta línea de herramientas para controlar la fiabilidad queda vinculada a los tests que emplean los psicómetras en sus laboratorios, es decir baterías de preguntas o pruebas de formato igual. No es aplicable a ninguna de las otras herramientas que se usan en un sondeo, y menos aún —obviamente— a todas las herramientas para recoger informaciones sobre unidades no individuales (agregados territoriales, etcétera).[39]

Además, la idea misma de que se pueda asignar a un test una etiqueta numérica que atestigüe para siempre su nivel de fiabilidad, como si fuera un atributo permanente y definitivo, sufre de una falla gnoseológica aún más grave que los varios límites técnicos que hemos visto arriba. Más que de una falla simple, se trata de un complejo de fallas que conciernen el problema de la generalizabilidad.

La más obvia es que no es legítimo generalizar a todas las poblaciones en todos los tiempos y situaciones un resultado obtenido con una población específica (habitualmente estudiantes de psicología) en un tiempo y situación específicos (habitualmente la clase del autor del test, y apenas éste ha sido imaginado). Esto parece obvio, pero "los coeficientes de fiabilidad encontrados en un grupo son transferidos a todos los otros grupos sin cuestionamientos" (Tukey 1954: 40). Se encuentran frecuentemente compilaciones de tests sobre varios temas donde cada test va acompañado por su coeficiente de fiabilidad;[40] muy raramente se reportan los resultados de controles sucesivos a la primera determinación del coeficiente.[41]

Una observación más sutil es que la generalización es ilegítima incluso entre el mismo grupo de sujetos a los que se aplica el test para calcular su coeficiente de fiabilidad, en cuanto:

a) cada sujeto tiene una manera particular de reaccionar a cada pregunta o prueba de un test;

[38] Una exhaustiva crítica técnica de todos los coeficientes de fiabilidad de la segunda generación está en Giampaglia (1986).

[39] Una observación análoga en Corbetta (2003: 108).

[40] Véanse, por ejemplo, Fiske (1971), Cronbach *et al.* (1972), Robinson *et al.* (1991).

[41] Una excepción es el manual de Shaw y Wright (1967), que reporta variaciones de un coeficiente .98 inicial a un .71 en el control sucesivo.

b) cada pregunta, o prueba, de un test tiene su particular nivel de inteligibilidad, de aceptabilidad, de agrado;
c) la reacción de cada sujeto particular a cada pregunta, o prueba, particular no es estable, sino que puede variar en cada situación específica.

Reduciendo toda esta variabilidad a un simple coeficiente que se pretende universal y definitivo se echa por la ventana la especificidad y no fungibilidad de los seres humanos, que había sido de alguna forma (aunque mínimamente) reconocida aplicando los tests a una pluralidad de individuos en lugar de sentar todas las teorías sobre un individuo aislado.

Como ha observado el psicólogo Frank:

> "Las técnicas universalmente usadas en psicología sólo dan informaciones acerca de un complejo de individuos, y nada sobre cada sujeto particular [...]. Estos tests, imaginados cuando no había otra manera de controlar la fiabilidad, ahora se han convertido en fetiches y bloquean la búsqueda de otras formas de control" (1939: 400).

Una línea crítica diferente ha sido expresada por el antropólogo Naroll:

> "El término 'fiabilidad' en su significado técnico de congruencia entre dos o más vectores no es atinado. El significado no técnico de 'fiable' es 'digno de confianza'. No creo que los científicos sociales vayan a tener más confianza en una herramienta por efecto de un test de fiabilidad. Se deberían llamar 'tests de congruencia' —lo que efectivamente son— más que 'tests de fiabilidad' —lo que en el sentido ordinario del término no son para nada" (1968: 265-6).

A varias décadas de distancia de estos escritos, la situación permanece problemática. Los controles basados en los coeficientes descritos en este apartado se han vuelto todavía más técnicos y alejados de todo contacto con situaciones reales. Para controlar realmente la fidelidad, sólo tenemos los procedimientos que se han descrito (apartados 7.1 y 7.3). Algunos (*wild code check*, *consistency check*, inspección de documentos oficiales) se pueden aplicar sólo raramente; otros (la entrevista sobre la entrevista, y formas parecidas) demandan una enorme inversión de tiempo, recursos y habilidad. Aparte de esto, no hay otro camino para mejorar la fidelidad de los datos que una gran atención al formular las definiciones operativas, soportada por una gran experiencia de todas las situaciones de peligro (para la fiabilidad) que se producen a lo largo del proceso de recolección de datos.

Algunas sugerencias serán dadas en los capítulos siguientes.

7.5. ¿Las definiciones operativas registran o manipulan opiniones y actitudes?

En el apartado precedente hemos visto que, limitándose al análisis de datos en la matriz y olvidando mirar afuera en el mundo, se pueden lograr coeficien-

tes de fiabilidad tranquilizadores que, sin embargo, no tienen ninguna relación con las situaciones reales.

Naturalmente, la responsabilidad de esta manera de proceder no es propia de la matriz, sino de la actitud de quienes la usan. La matriz en sí es un instrumento neutro, que se puede usar también para descubrir indicaciones significativas sobre dichas situaciones reales. Lo hizo hace medio siglo Donald T. Campbell, el metodólogo más dotado de espíritu crítico y de sensibilidad gnoseológica.[42] Con la ayuda del psicólogo Donald W. Fiske pasó revista a los millares de investigaciones depositadas en los archivos de la Universidad de Michigan, con el fin de identificar todas las que habían registrado propiedades psíquicas (opiniones, actitudes, valores) con tres o más definiciones operativas diferentes. Para cada una de las investigaciones que cumplían con estos requisitos Campbell y Fiske construyeron una matriz del tipo que se ve en la tabla 7.1, que bautizaron *multitrait multimethod matrix* (matriz con muchas propiedades y muchas técnicas).

Tabla 7.1: *Multitrait multimethod matrix*

	Al	Aa	At	El	Ea	Et	Pl	Pa	Pt
Al	1	p	p	*t*			*t*		
Aa	p	1	p		*t*			*t*	
At	p	p	1			*t*			*t*
El	*t*			1	p	p	*t*		
Ea		*t*		p	1	p		*t*	
Et			*t*	p	p	1			*t*
Pl	*t*			*t*			1	p	p
Pa		*t*			*t*		p	1	p
Pt			*t*			*t*	p	p	1

En el ejemplo de la tabla, se consideran tres propiedades psíquicas (autoritarismo, extraversión, particularismo) cada una registrada con tres diferentes técnicas (escala de Likert, escala autoanclante, termómetro: véase apartado 8.5). La tabla representa una matriz de correlación (véase apartado 15.6) Cada fila representa una propiedad particular registrada con una técnica particular: así Al significa autoritarismo registrado con una escala de Likert, Aa significa autoritarismo registrado con una escala autoanclante, etcétera. Como en todas las matrices de correlación, la secuencia de elementos en las filas se repite idéntica en las columnas, y en cada celda está el coeficiente de correlación entre la variable en la fila correspondiente y la variable en la columna correspondiente. Naturalmente, en las celdas de la diagonal descendente están las correlaciones de una variable consigo misma, que tienen el valor 1. En la tabla, las celdas señaladas con una p contienen la correlación entre variables que registran la misma propiedad con técnicas diferentes; las celdas señaladas con una *t* contienen la correlación entre variables que registran propiedades diferentes con la misma técnica; las celdas blancas contienen la correlación entre variables que registran propiedades diferentes con técnicas diferentes.

42 Se trata del mismo que en obras clásicas (Campbell y Stanley 1963; Cook y Campbell 1979) criticó la simpleza de muchos "experimentos" en las ciencias sociales, estableciendo una serie de condiciones y de procedimientos que permiten reducir la gran distancia epistemológica entre los experimentos de las ciencias físicas y lo que se puede lograr en las ciencias sociales. Véase apartado 1.2.

Si la técnica de registro no tuviese ningún impacto, es decir no influyese mínimamente en la forma en que los estados en las propiedades son convertidos en datos, no habría razón para esperar que las correlaciones entre variables que registran propiedades diferentes con la misma técnica fuesen más altas que las correlaciones entre variables que registran propiedades diferentes con técnicas diferentes; las únicas correlaciones fuertes se deberían encontrar entre las variables que registran la misma propiedad con técnicas diferentes. Usando los símbolos de las celdas de la tabla 7.1, la expectativa sería:

$$\boxed{p} > \boxed{t} = \boxed{}$$

Por el contrario, Campbell y Fiske encontraron que en la gran mayoría de las matrices *multitrait multimethod* construidas con los datos de las investigaciones depositadas en los archivos, la situación efectiva era:

$$\boxed{t} \gg \boxed{p} > \boxed{}$$

Es decir, los coeficientes de correlación entre las variables eran casi siempre más o mucho más altos cuando sólo la técnica era la misma que cuando sólo la propiedad era la misma. En otras palabras, las técnicas determinaban los estados que sólo deberían haber registrado.

Cuando Campbell y Fiske divulgaron sus resultados (1959), el artículo tuvo el impacto de un sismo en la comunidad de investigadores que adoptaban la aproximación estándar. Durante años no se publicó un solo informe de investigación sobre opiniones y actitudes en el que no se apelase ritualmente a él (y luego se siguiese empleando las mismas técnicas que el artículo ponía en tela de juicio). Lentamente, incluso los homenajes formales se hicieron raros, y hoy en día el artículo se menciona sólo como un destacado antecedente de la tendencia actual hacia la triangulación metodológica (véase apartado 2.4).

Capítulo 8

CLASIFICACIÓN, CONTEO, MEDICIÓN, CONSTRUCCIÓN DE ESCALAS

Hasta ahora hemos considerado las propiedades (y por lo tanto las variables que las representan en la matriz de datos) globalmente. En este capítulo consideramos los diversos estados en cada propiedad, las consecuencias de la naturaleza de esos estados sobre la manera en que se transforman en datos y el tipo de análisis al cual se pueden someter las variables correspondientes.

Algunas propiedades, tanto de objetos físicos como de individuos o de unidades de otro tipo, tienen un número finito de estados, claramente distintos (o distinguibles) el uno del otro: en una palabra, discretos. Y DISCRETAS se llaman las propiedades que tienen estados discretos.

Las otras propiedades tienen un número infinito de estados imperceptiblemente diferentes el uno del otro: en este caso se habla de propiedades CONTINUAS.

Veremos pronto ejemplos. Por el momento, cabe destacar que la distinción entre propiedades discretas y propiedades continuas es la más importante cuando se trata de recoger informaciones y transformarlas en datos por medio de una definición operativa.

Las propiedades discretas se dividen en tres clases según el tipo de relación que existe entre sus estados.

Si los estados son sólo diferentes el uno del otro, pero no hay ningún tipo de relación cuantitativa entre ellos, se habla de propiedades CATEGORIALES. Ejemplos: la nacionalidad de un individuo, el tipo de régimen político de un estado, el tipo de producción agrícola prevaleciente en un territorio, etcétera (véase apartado 8.1).

Si los estados tienen relaciones ordinales (de mayor y menor) entre ellos, es decir, se pueden ordenar a lo largo de una escala, se habla de propiedades ORDINALES. Ejemplos: el nivel de estudio alcanzado por un individuo, su grado jerárquico en una organización civil o militar; el grado de centralidad de una comuna en su provincia, el grado de dureza de un metal (escala de Mohs), etcétera (véase apartado 8.2).

Si los estados tienen relaciones cardinales, en el sentido de que se puede legítimamente establecer un cociente entre ellos, se habla de propiedades DISCRETAS CARDINALES. Ejemplos: el número de componentes de una familia, el nú-

mero de tarjetas que tiene un individuo, el número de camas en un albergue, el número de licenciados en una provincia, etcétera (véase apartado 8.3).

Las propiedades continuas se dividen en dos clases. Si sus estados se pueden registrar sin intervención activa del individuo cuyo estado se está registrando, se habla de propiedades MENSURABLES (véase apartado 8.4). Si esta intervención activa es necesaria se habla de propiedades CONTINUAS NO MENSURABLES (véase apartado 8.5).

8.1. Propiedades y variables categoriales

Consideremos los tres ejemplos de propiedades categoriales: en uno de ellos los estados en la propiedad (nacionalidad de un individuo) son claramente distintos; en los otros dos (tipo de régimen político, tipo de producción agrícola) se necesita una intervención de expertos para abordar casos dudosos, intermedios, mixtos, etcétera.

En la realidad, hay una gran variedad de estados en las propiedades (piénsese en cuántas diferentes producciones agrícolas pueden existir en los campos de un territorio); se debe decidir qué estados se convertirán en categorías de la variable que va a representar la propiedad en la matriz.

Esta intervención es necesaria incluso cuando la transformación de estados en categorías parece simple y automática. Por ejemplo, en el caso de la nacionalidad, se deben tomar al menos dos tipos de decisiones:

- cómo tratar a los individuos que tienen dos nacionalidades, o más (¿se crearán categorías mixtas o se tomará el país donde nació el individuo o el país donde reside habitualmente?)
- si y cómo agregar las nacionalidades con pocos individuos en la muestra (¿se agregarán las nacionalidades de algunos continentes, por ejemplo "africanos", o de algunas áreas geográficas, por ejemplo "caribeños" o se seguirá cualquier otro criterio que sugieren los objetivos de la investigación?)

Si una intervención del investigador es necesaria en estos casos, que parecerían simples, tanto más lo es cuando se trata de propiedades —como opiniones, actitudes, valores— que no tienen estados socialmente codificados, y cuyos estados tienen que ser inferidos, basándose en declaraciones o actos del sujeto, antes de convertirlos en categorías de la variable.

Todas estas operaciones de creación y/o delimitación de categorías se realizan siguiendo los principios de una operación intelectual —la CLASIFICACIÓN— que tiene sus raíces en el pensamiento griego.[1] Esta operación se puede realizar

[1] También en la India clásica hay doctrinas acerca de la clasificación, formuladas de manera más orgánica en el *Vaisesika sutra* (siglo I d.C.).

en abstracto,[2] considerando las diversas formas que puede tomar un objeto o evento (en nuestro caso, los múltiples estados que se pueden concebir en la propiedad) o bien observando los diversos ejemplares de un objeto o evento (en nuestro caso, las múltiples manifestaciones de una propiedad) y agrupando las que nos parecen semejantes.[3]

El criterio que se sigue para dividir la intensión de un concepto o para agrupar los ejemplares tiene un nombre de origen tardo-latino:[4] *fundamentum divisionis*. Por ejemplo, si se divide un conjunto de individuos en morochos, rubios, pelirrojos y canosos, estamos adoptando el color del pelo como *fundamentum divisionis*.

Si se adoptan conjuntamente dos o más criterios (por ejemplo, si se divide el grupo en hombres morochos, mujeres morochas, hombres rubios, etcétera) la operación intelectual se llama TIPOLOGÍA. Algunas veces, un solo *fundamentum divisionis* se considera insuficiente para establecer las categorías de una variable, y por lo tanto se usan dos o más, creando una tipología.[5]

Lo importante es que las categorías establecidas sean mutuamente excluyentes: es decir que un ejemplar (o un estado) pertenezca —o sea asignado— a una sola categoría.

Este requisito de la MUTUA EXCLUSIVIDAD vale tanto para las tipologías como para las clasificaciones; sin embargo, es frecuentemente violado, por ejemplo, por los medios masivos de comunicación. Un caso claro pasó durante las guerras en Yugoslavia, cuando los medios europeos hablaban de serbios, croatas y musulmanes, mezclando[6] dos criterios (idioma y religión) y por lo tanto violando la mutua exclusividad, en cuanto la mayoría de los musulmanes del área eran (y son) de habla serbia o croata.

Respetar el requisito de mutua exclusividad de las categorías es fácil cuando se clasifican objetos abstractos, como por ejemplo los números.[7]

Otra cosa es respetarlo en la investigación social, cuando por ejemplo se deben clasificar las respuestas espontáneas a preguntas "abiertas",[8] o las reacciones a un test proyectivo,[9] o los significados de un discurso o de un docu-

[2] En este caso se habla de clasificación intensional, ya que el complejo de los aspectos de un concepto se llama 'intensión'. Esta aproximación es típica de Platón y de su escuela.

[3] En este caso se habla de clasificación extensional, ya que se llama 'extensión' al complejo de objetos/eventos que cumplen los requisitos para ser considerados ejemplares de un concepto. Esta aproximación es típica de Aristóteles y su escuela, y constituye el fundamento de las modernas técnicas de *cluster análisis* (análisis de conglomerados: véase Aldenderfer y Blashfield 1985).

[4] Los filósofos escolásticos del Medioevo, que escribían en latín, estudiaron con gran sutileza los problemas vinculados a la clasificación.

[5] En este caso, las categorías no se llamarán clases, sino tipos.

[6] Sólo combinando, en cambio de mezclar, los criterios se conseguirá una tipología correcta: croatas cristianos, croatas musulmanes, serbios cristianos, serbios musulmanes.

[7] No cabe duda en establecer cuáles son los números pares y cuáles los impares, los enteros y los decimales, los primos y los múltiplos, etcétera.

[8] Una pregunta se llama abierta cuando se deja que el entrevistado conteste libremente en vez de invitarlo a elegir una respuesta entre un abanico de respuestas preestablecidas.

[9] Se llaman proyectivos los tests en que se invita al sujeto a dar su interpretación de una figura ambigua (test TAT: Murray 1938), e incluso de una mancha de tinta (Rorschach 1921). Habitualmente se llaman proyectivos también los tests en que se invita al sujeto a completar una corta frase inconclusa (Jung 1910) o a reaccionar a una agresión verbal dada (Rosenzweig 1938).

mento oficial. Al ejecutar estas tareas es fácil toparse con situaciones en que el estado de un caso en una propiedad (es decir sus respuestas a una pregunta, o reacciones a un dibujo, etcétera) sea legítimamente asignable a dos o más categorías. Para garantizar el máximo posible de homogeneidad en estas asignaciones, quien(es) establece(n) la definición operativa tiene(n) que preocuparse de complementar las reglas con una rica casuística de situaciones problemáticas acompañadas por las soluciones correspondientes.

Otro requisito clásico de la clasificación es la EXHAUSTIVIDAD de las categorías: todos los estados posibles en una propiedad deben poder ser asignados a una categoría.[10]

Ya que a menudo no es posible, o no sirve, imaginar todos los estados posibles en una propiedad, muchas clasificaciones prevén una categoría "otro", a la cual se asignan todos los estados que no parecen asignables a las categorías que tienen un significado propio. La categoría "otro", en cambio, no tiene ningún significado, ya que cumple el rol residual, donde finalizan todos los estados no previstos.

Cabe distinguir la categoría "otro" de categorías que podrían parecer igualmente residuales, como "no contesta", "no sabe", habitualmente usadas para clasificar respuestas a preguntas planteadas en un sondeo. A menudo estos "datos faltantes" son descuidados en el análisis, mientras que pueden dar informaciones útiles, ya sea sobre los entrevistados, ya sea sobre el cuestionario que se está usando.

Una cuota demasiado alta de "no sabe" puede significar que algunos temas del cuestionario no interesan a la mayoría de nuestros entrevistados, o que las preguntas han sido formuladas con un lenguaje alejado del lenguaje ordinario que ellos hablan y entienden.

Una cuota demasiado alta de "no contesta" puede significar que algunas preguntas tocan temas que la mayoría de los entrevistados juzga delicados y privados.

Por otro lado, una cuota demasiado alta de estados que se deben asignar a la categoría "otro" significa que la clasificación fue realizada de manera superficial y/o sin una suficiente familiaridad con la población y con el tema que se investiga.

Estas faltas (preguntas planteadas sobre temas delicados, poco interesantes, poco conocidos por los sujetos o por los investigadores, etcétera) ocurren en muchos sondeos, y se podrían evitar con una PRUEBA PILOTO, es decir, sometiendo preliminarmente el cuestionario a algunos miembros, lo más diferentes entre sí que sea posible, de la misma población que se pretende investigar.

Reflexionando adecuadamente sobre los resultados de esta prueba piloto se pueden introducir mejoras en el texto del cuestionario. Por el contrario, casi siempre se ejecutan estas pruebas piloto de forma ritual (para quitarse de encima este deber convencionalmente reconocido) e inmediatamente antes del propio sondeo, de forma que no hay tiempo para analizar los resultados y, en efecto, éstos se guardan sin dedicarles atención.

[10] La mutua exclusividad es un requisito de cada pareja de clases (o tipos); la exhaustividad es un requisito de cada conjunto de clases. Consideradas conjuntamente, ellas aseguran que cada ejemplar (en nuestro caso, cada estado) sea asignado a una, y una sola, clase (o tipo).

Al considerar el problema de la exhaustividad, hemos gradualmente pasado de los requisitos clásicos de clasificaciones y tipologías a algunos problemas prácticos relacionados con la investigación social. Veamos ahora algunos otros problemas.

El primero es la elección del *fundamentum divisionis*, es decir del criterio según el cual se establecen las categorías. Obviamente se elegirá un *fundamentum* relevante para nuestros intereses: por ejemplo, si se debe elegir el criterio para clasificar una clase de jóvenes al inicio de un curso, un *fundamentum* como el tipo de estudios precedentes nos ayudará a decidir el nivel y la orientación de éste.

El problema cognoscitivo que se pretende profundizar con una investigación proporciona indicaciones importantes, no sólo sobre los *fundamenta* de cada clasificación que se realizará, sino también sobre su sensibilidad.[11] Un estadístico que esté colaborando con el gobierno indio para el censo arreglará una clasificación extremadamente sensible de la propiedad "religión", porque sabe que la Constitución india atribuye derechos y tutelas particulares a los seguidores de algunas religiones de minoría. Por el contrario, un sociólogo que programa un sondeo telefónico preelectoral con mil entrevistados y quiere plantear una pregunta sobre la religión porque ella puede influir en el voto, sería torpe si previera categorías específicas para cada secta menor: si un seguidor de una de esas sectas posee un teléfono y es extraído en la muestra, el destino de su estado en la propiedad "religión" será la categoría "otro".

Con la excepción del sondeo telefónico —por motivos que veremos pronto—, se puede establecer una regla general: cabe recolectar informaciones lo más detalladas posible, y por lo tanto se adoptará una clasificación muy sensible. Esto por dos motivos:

1) como se dijo en el apartado 3.3, cuando se atribuyan dos estados a la misma categoría, y por lo tanto se les asigne el mismo código en la variable, los dos estados serán considerados perfectamente iguales en cualquier análisis sucesivo;
2) es siempre posible y fácil reunir dos o más categorías de una variable antes de analizar sus relaciones con otras variables, y muchas veces esto es fuertemente aconsejable, por motivos que veremos; pero, es casi siempre imposible[12] desagregar una categoría a la cual fueron asignados estados parecidos, en el caso de que en lo sucesivo se quiera distinguirlos.

[11] Sobre el concepto de sensibilidad, véase el apartado 7.1.

[12] Esto es posible sólo si —por ejemplo— en el mismo cuestionario hay otra pregunta que resulte adecuada a tal efecto. Un caso se verá más adelante en este apartado.

Sin embargo, esta regla no se debe aplicar en el caso de las preguntas "cerradas" de un sondeo, en el cual el entrevistador plantea oralmente al sujeto una gama de respuestas predefinidas para que él elija una. En este caso no se pueden presentar demasiadas respuestas, porque nuestra capacidad de considerar simultáneamente varias alternativas es muy limitada.[13]

El entrevistado medio tiende a recordar siempre la última alternativa oída,[14] a menudo también la primera, y casi nunca las intermedias, lo que conlleva que cada individuo considere un número reducido, variable y —sobre todo— desconocido de alternativas en lugar de considerar todas las que se le ofrecen. Este sesgo es aún más grave en la entrevista telefónica, donde la interacción es habitualmente corta y apresurada, se dan problemas de recepción y de comprensión, y el entrevistado no goza de todas las señales, voluntarias o no, que recibiría con la presencia física del entrevistador.[15]

Concluyendo sobre este problema de sensibilidad, se aconseja el máximo detalle cada vez que las informaciones sean recogidas sin interacción oral con el sujeto. Un detalle mucho menor (pocas categorías generales y claramente distintas) si el sujeto tiene que escuchar las alternativas y elegir una sin poderse apoyar en un texto escrito. Si la entrevista es telefónica, se aconseja limitarse a dicotomías (sí o no, a o b) o a propiedades ordinales sencillas.[16]

Cuando se establezcan las categorías, se tratará de asignar un código a cada una. Con la asignación de códigos se cumple el paso más importante en el proceso de la transformación de la propiedad categorial en una variable categorial. Puesto que las categorías de este tipo de propiedad no tienen ninguna diferencia cuantitativa entre sí, no hay ninguna razón para que estos códigos sean numéricos.[17] Con todo, ésta es una mala costumbre generalizada, que conlleva el riesgo de que algunos traten estos aparentes números como si fueran números y sometan variables categoriales a técnicas de análisis que en sentido estricto, sólo se pueden aplicar a variables cardinales, produciendo ridículos sinsentidos.[18] Este riesgo se ha acrecentado por la difusión universal de la expresión 'escala nominal' introducida por Stevens (1946), que tiene el efecto perverso de justificar el empleo de técnicas cuantitativas con estas pretendidas escalas.[19]

[13] En un afamado artículo, el gran psicólogo cognitivista George Miller (1956), basándose en investigaciones de laboratorio, puso este límite en siete alternativas. Pero en una típica entrevista "cara a cara", puertas afuera del laboratorio, parece que el límite resuta más bajo. Johann Galtung (1967) lo pone en torno de cuatro o cinco alternativas; Sudman y Bradburn (1982) en no más que cinco. En la experiencia de los autores de este libro, si las alternativas no son sencillas, muchos entrevistados tienen dificultades en considerar más de dos. Más en general, acerca de la discusión sobre preguntas "abiertas" o "cerradas", véase, por ejemplo, Converse (1984).

[14] En la literatura metodológica se habla en este caso de *end effect* (efecto fin). El fuerte impacto de este *end effect* ha sido demostrado confrontando las distribuciones de las respuestas elegidas por muestras diferentes de individuos a los que se presentaron las mismas alternativas cambiando su orden. Por otro lado, si el entrevistado lee las alternativas en lugar de oírlas, su atención se concentra en las primeras de la lista (Corbetta 2003, 185).

[15] Sobre la entrevista telefónica y la telemática, véase Corbetta (2003, 195-202).

[16] En el apartado siguiente se verá por qué propiedades y variables ordinales presentan menores problemas de discriminación y valoración.

[17] Los códigos numéricos introducen inevitablemente al menos la idea de un orden entre categorías. Pero no hay ningún orden entre, por ejemplo, las nacionalidades, y por lo tanto la decisión de asignar el código (1) a mexicanos y el código (2) a peruanos en lugar del contrario es totalmente arbitrario.

Sea cual sea (numérico o no) el código que se le asigna, cuando tenga un código y esté siendo analizada con técnicas estadísticas, una categoría cambia de nombre y se llama 'modalidad'. Con la expresión 'frecuencia de una modalidad' se entiende el número de datos a los cuales se atribuyó aquella modalidad. La modalidad con la frecuencia más alta se llama MODA (o modo) —y esto en todos los tipos de variable.

El primer paso del análisis de una variable categorial es la producción de su DISTRIBUCIÓN DE FRECUENCIA, es decir de una tabla que tiene habitualmente cuatro columnas: en la primera se pone el código (numérico o no), en la segunda se pone el nombre de la categoría/modalidad, en la tercera la frecuencia (absoluta) y en la cuarta el porcentaje correspondiente (frecuencia relativa).

Tabla 8.1: *Una distribución de frecuencia de la variable "nacionalidad"*

CÓDIGO	CATEGORÍA/MODALIDAD	FRECUENCIA	PORCENTAJE
a	argentinos	72	21,5
b	brasileños	126	37,5
c	chilenos	34	10
e	ecuatorianos	15	4,5
p	peruanos	40	12
u	uruguayos	26	8
0	otros	22	6,5
Total		335	100

Tres observaciones:

1) como códigos no se pusieron números, sino la primera letra del nombre de las categorías;
2) para distinguir la categoría residual (otros), que como se dijo no tiene sentido unitario porque agrega cualquier nacionalidad no explícita-

[18] El sinsentido depende del hecho de que —como se subrayaba en la nota previa— la asignación de números a categorías que no tienen ninguna característica cuantitativa no puede ser totalmente aleatoria. Cambiando esta asignación, también cambia el resultado del análisis. Un ejemplo clamoroso de sinsentido lo proporciona el conocido metodólogo norteamericano Rudolph Rummel, que ha insertado en una ecuación de regresión múltiple las siguientes variables categoriales: tipo de sistema electoral, naturaleza del régimen, tipo de liderazgo político, tipo de intervención militar en la política, calculando coeficientes cuyo valor depende totalmente de una decisión arbitraria sobre los códigos numéricos asignados a las categorías (1963: 203).

[19] La expresión fue introducida, en un período de conductismo radical, en lugar de términos como 'clasificación' y 'categoría' que tenían una tradición filosófica secular y por lo tanto un sonido poco agradable a oídos cientificistas. Pero el adjetivo 'nominal' (las categorías se diferencian sólo por el nombre) contradice al sustantivo 'escala', vinculado a una idea de cantidad. En casos como éste los escolásticos hablaban de *contradictio in adjecto* (el adjetivo contradice al sustantivo).

mente considerada en la lista, se le puso el código numérico cero, que da la idea de nulidad.

3) en general, los porcentajes no tienen decimales porque las frecuencias son relativamente bajas. En este caso, cabe redondear los decimales al entero más cercano. Una excepción (que se observa también en la tabla 8.1) puede ser el decimal 5, a medio camino entre los dos enteros: éste se puede usar para evitar redondeos demasiado fuertes.[20]

Por el contrario, a menudo se ven tablas como la siguiente, donde se dan porcentajes con dos cifras decimales calculados sobre un total de 45 sujetos.[21]

Tabla 8.2: *Un caso de "falacia de la precisión fuera de lugar"*

CATEGORÍA/MODALIDAD	FRECUENCIA	PORCENTAJE
burócratas	12	26,67
desempleados	8	17,77
educadores	7	15,55
obreros	6	13,34
comerciantes	6	13,34
jubilados	3	6,67
artesanos	2	4,45
otros	1	2,22

Fuente: Saviane (1978).

Esta conducta ridícula es tan frecuente que ha recibido un nombre (*fallacy of the misplaced precision* = falacia de la precisión fuera de lugar). Su motivo ha sido agudamente detectado por el metodólogo Abraham Kaplan:

> "A menudo se encuentran ejemplos de una precisión sin sentido y sin motivación, o de una precisión más grande de la que nuestras herramientas permiten [...] El anhelo que muchos sociólogos tienen de exhibir esta precisión insensata refleja un sentido de inseguridad y un complejo de inferioridad con respecto a las ciencias físicas" (1964: 204-5).

Y se puede compartir el comentario de Jerome Ravetz: "Nada manifiesta la falta de cultura matemática como la falacia de la precisión fuera de lugar" (1971: 45).

[20] Una tabla práctica de redondeos puede ser la siguiente:

10,1 10,2 10,3	10	10,4 10,5 10,6	10,5	10,7 10,8 10,9	11

[21] Otros ejemplos: Sebastani (1970) da porcentajes con dos decimales calculados sobre un total de 60 sujetos; Stern (1975) da porcentajes con un decimal calculado sobre un total de 14 sujetos.

Mientras que a veces los porcentajes son ridículamente detallados, otras veces pasa que en la tabla las frecuencias absolutas faltan totalmente y sólo se presentan los porcentajes. Esta laguna puede tener un solo motivo: la muestra fue tan pequeña que no se quiere dar información (las frecuencias) que revelaría esta grave falta de la investigación.

> Si la muestra es pequeña, algunas categorías pueden tener frecuencias de 1 o de 2. Si es así, todos los porcentajes presentados en la tabla deben, por necesidad matemática, ser múltiplos del porcentaje más bajo o de su mitad. Si controlando (y teniendo en cuenta los redondeos) el lector encuentra que todos los porcentajes son múltiplos, no tendría dificultad en calcular el total de los casos de aquella tabla.

La inspección visual de sus distribuciones de frecuencia es la forma preliminar de análisis de variables categoriales. Ella permite tomar contacto con los datos que hemos recolectado. Además, puede señalar la presencia de errores en la codificación, si en la distribución aparecen códigos que no se habían previsto en nuestro plan (*wild codes*: véase apartado 7.3). Ella señala también desequilibrios en las frecuencias (modalidades con un número demasiado alto o demasiado bajo), que pueden también indicar errores en la codificación y que de todas formas provocan problemas en algunos tipos de análisis estadístico (véase más adelante).

Una vez realizado este paso preliminar, una variable categorial está casi lista para ser puesta en relación con otras variables categoriales u ordinales mediante una tabla bivariable, o con variables cardinales mediante el análisis de la varianza. Todo esto se verá en el capítulo 15.

Se dijo "casi lista" y no "lista" porque si se quiere investigar de manera eficaz las relaciones de las variables categoriales con otras variables, hay que someter casi todas a una especie de "maquillaje", que constituye una verdadera fase intermedia entre recolección y análisis. Aunque algunos metodólogos expertos dedican atención a esta fase, ella es habitualmente descuidada en los textos de metodología.[22]

Se dijo que al recolectar informaciones hay que privilegiar el detalle, y por lo tanto la sensibilidad de nuestras clasificaciones. Hay que reproducir el máximo de detalle también cuando los datos recolectados se exploran a través de una distribución de frecuencia. Sólo con una representación detallada podemos tener un conocimiento profundo de las situaciones que hemos explorado. Pero cuando se quieran analizar las relaciones de una variable categorial con otras variables, el discurso cambia radicalmente: el mismo nivel de detalle que es positivo en la recolección y análisis univariado se convierte en negativo.

El grado de esta negatividad depende de la naturaleza de la otra variable con la cual establecemos la relación: si es una cardinal, el problema no es grave, como veremos; si es una ordinal, éste resulta mayor, pero aún no tan grave; si

[22] Como excepción señalamos por ejemplo el ensayo de Agnoli (1992).

es una categorial, un alto nivel de detalle —es decir, un gran número de categorías— puede producir dificultades importantes. Esto porque en la herramienta con la que se analizan las relaciones entre dos variables categoriales —la tabla bivariable— el número de las celdas equivale al producto de las modalidades de las dos variables que forman la tabla. Con dos variables de diez modalidades, el número de celdas en la tabla es cien. Con dos variables de veinte modalidades, el número es cuatrocientos. Esto crea dos problemas: uno meramente numérico, el otro semántico.

Veamos el primero. Cada sujeto de la muestra tiene su combinación de estados en las dos propiedades (y por consecuencia, de códigos en las dos variables). Mil sujetos, mil combinaciones (obviamente, algunas serán iguales). En cada celda de la tabla se encuentra la combinación de una modalidad de la variable que está en las filas y una modalidad de la variable que está en las columnas. Si las celdas de la tabla son cien y los casos en la matriz son mil, evidentemente el promedio de las frecuencias (número de datos) en las celdas será diez. Con 400 celdas, el promedio será 2,5.

Obviamente, algunas combinaciones serán más frecuentes que el promedio y otras menos —lo que significa, con promedios de 10 o menos, que algunas celdas estarán vacías o casi vacías. De celdas vacías o casi vacías no se pueden sacar conclusiones creíbles acerca de las relaciones entre las modalidades que forman esta celda: por eso una tabla con muchas celdas vacías o casi vacías no sirve mucho para el análisis.[23]

Éste es el problema numérico. Para ilustrar el problema semántico hay que introducir el concepto de AUTONOMÍA SEMÁNTICA. Una categoría tiene plena autonomía semántica si puede ser interpretada sin hacer referencia al significado de la propiedad o de las otras categorías. Por su naturaleza, las categorías de las variables categoriales tienen plena autonomía semántica. Lo que significa que cada combinación de las categorías de dos variables categoriales (y por consecuencia cada combinación de modalidades en la tabla) tiene su sentido y necesita atención y concentración para enfocar su pleno significado. Para juzgar si una frecuencia particular en cada celda es o no es interesante hay que concentrarse en lo que significan las dos modalidades que la forman, y en lo que significa su relación: una operación intelectual que con una normalísima tabla de cien celdas puede tomar horas, y con una de 400 celdas podría tomar un día. Puesto que la computadora permite producir centenares de tablas en segundos, el resultado de todo esto son montañas de papel que nadie tiene el tiempo para analizar.

En vez de producir montañas de papel, el investigador experto prefiere enfocar su atención en pocas tablas elegidas con cuidado y hacerlas manejables, reduciendo al mínimo el número de celdas vacías o casi vacías, de manera de

[23] La razón de esto se entenderá mucho mejor tras leer el apartado 15.2, en el cual se introducen los conceptos de frecuencia esperada y frecuencia empírica (esta última se suele denominar "frecuencia observada"; pero en este texto se prefiere decir "empírica", porque en las ciencias sociales raramente se observa en sentido estricto).

poder aprehender todo lo que indica la tabla en un tiempo razonable. Para lograr este objetivo hay un solo camino: reducir el número de las categorías en ambas variables.

Esta operación se lleva a cabo de forma correcta y eficaz si se siguen dos criterios: uno medular y otro que a veces puede mitigarlo. El criterio medular es la proximidad semántica entre las categorías que se quiere unir. Naturalmente esta proximidad tiene que ser juzgada bajo la luz de los intereses cognoscitivos por los que se analiza la variable y sus relaciones con otras.

> En una clasificación de profesiones de individuos, se pueden unir terratenientes y empresarios industriales si lo que interesa es el tipo de relación con los medios de producción; se pueden unir terratenientes y campesinos si lo que interesa es el sector de actividad; no se pueden unir campesinos y empresarios industriales (a menos que se adopten extraños criterios de proximidad semántica).

El criterio auxiliar es el equilibrio entre las frecuencias en las distintas modalidades. Por razones que veremos pronto, el metodólogo experto trata de evitar la presencia de modalidades, y por lo tanto celdas, con frecuencias demasiado altas o demasiado bajas, ya que ambas perjudican la interpretabilidad de una tabla. El arte del metodólogo, de lo que se hablaba en el capítulo 3, consiste en el balanceo, en cada situación, del impacto de estos dos criterios (semántico y numérico). Retomando el problema del estadístico indio que exploraba la influencia de las religiones en el voto, si al recoger los datos él usó una clasificación con treinta categorías, sería sabio, a la hora del análisis, reagrupar las religiones menos difundidas en algunas categorías más amplias, poniendo en cada nueva categoría las que tienen una orientación teológica y filosófica similar.

> Veamos un caso particular, sacado de una investigación que en los años noventa uno de los autores de este libro llevó a cabo por encargo de la asociación de jóvenes empresarios italianos. El objetivo de la investigación era detectar las formas de impacto sobre las actividades empresariales que tenían las diversas organizaciones criminales.[24]
> Los 8.000 miembros de la asociación recibieron un cuestionario postal enviado por la misma asociación con garantía de anonimato. Una compleja pregunta pedía a cada joven empresario que juzgara cuál era, en cada sector económico, la fuente principal de capitales para las empresas: autofinanciamiento, préstamo bancario, subvenciones europeas, nacionales, locales, lavado de fondos de origen ilícito u otra. Las informaciones recogidas se podían organizar en forma de tabla poniendo los sectores en las filas y la fuente juzgada prevaleciente en las columnas. Puesto que los sectores productivos eran doce y las posibles fuentes siete, la tabla tenía 84 celdas. Pero observando atentamente las celdas se podía detectar un patrón: muchos

[24] Que tienen un nombre tradicional diferente en cada región italiana: mafia en Sicilia, camorra en Campania, ndrángheta en Calabria, sacra corona unita en Puglia, etcétera.

empresarios juzgaban que el lavado de dinero fuese la fuente principal en el sector de la construcción; ninguno juzgaba que lo fuese en los sectores de más alto contenido tecnológico (electrónico, metalmecánico, químico, etcétera); pocos juzgaban que lo fuese en los sectores de contenido tecnológico medio-bajo (textil, alimentario, etcétera). Al agrupar bajo este criterio a los doce sectores en sólo tres, el juicio de los jóvenes empresarios sobre el problema investigado emergía clarísimo de la tabla.

Con la salvedad de casos particulares como éste, donde el reagrupamiento se hizo considerando sólo el criterio semántico y descuidando el numérico, hay que tener en cuenta —en cuanto sea posible— el equilibrio entre las frecuencias de las categorías reagrupadas: el ideal sería un equilibrio perfecto (el mismo porcentaje en cada categoría).

Las consecuencias de graves desequilibrios son conocidas: como ya se indicó, categorías con frecuencias bajas producen inevitablemente celdas vacías o casi vacías y, como veremos en el apartado 15.3, las conclusiones acerca de sus relaciones con las categorías de otras variables no son confiables.[25]

A veces para lograr una distribución más equilibrada no basta el reagrupamiento de categorías semánticamente parecidas y con escasas frecuencias; se debería también desagregar una categoría que resultó tener un porcentaje excesivo. Pero esto —como se dijo— sólo es posible si en la misma matriz existe otra variable que puede ser empleada para desagregar aquella categoría.

Se puede tomar un ejemplo de un análisis secundario que uno de los autores de este libro realizó sobre datos recogidos en una investigación sobre 7.000 jóvenes italianos en los setenta. El municipio de residencia del joven era clasificado en una escala de seis niveles (urbanos, casi urbanos, semiurbanos, semirrurales, casi rurales, rurales). El 52% de los jóvenes entrevistados resultaron "urbanos". Las frecuencias de las otras categorías eran como se ve en la segunda columna de la tabla 8.3.

Tabla 8.3: *Ejemplo de agregación y desagregación de categorías*

CATEGORÍAS ORIGINARIAS	%	CATEGORÍAS FINALES	%
urbanos	52	urbanos céntricos	32
cuasi urbanos	4	urbanos periféricos	17
semiurbanos	5	semiurbanos	9
semirrurales	8	semirrurales	17
cuasi rurales	6	rurales	19
rurales	19		
inform. faltantes	6	inform. faltantes	6
número de casos	7.024		7.024

[25] Además, desequilibrios en la distribución producen graves sesgos en todos los coeficientes que cuantifican las relaciones entre variables, y esto *sea cual sea la naturaleza de las variables* (categorial, ordinal o cardinal).

Sólo la categoría "rurales" tenía una frecuencia cercana a la ideal, y por lo tanto no se modificó. El primer paso fue reagrupar las categorías intermedias: los casi urbanos se unieron a los semiurbanos, y los casi rurales a los semirrurales. Para desagregar la categoría "urbanos" se usó otra pregunta del cuestionario, con la que se preguntaba a los jóvenes clasificados "urbanos" si vivían en el centro, en la periferia o en el campo alrededor de la ciudad. Basándose en esta información se dividió a los jóvenes urbanos en dos categorías (urbanos céntricos y urbanos periféricos). Quedaba un 3% de jóvenes que habitaban en el campo alrededor de una ciudad. Ellos fueron unidos a los semirrurales, con el resultado que se ve en la última columna de la tabla 8.3: de las categorías resultantes, tres tienen un porcentaje cercano al ideal —que ahora es del 20% ya que las categorías son cinco (obviamente en este cálculo no se considera la categoría de informaciones faltantes).

8.2. Propiedades y variables ordinales

Como se decía antes, son ordinales las propiedades en las que se pueden ordenar todos los estados a lo largo de una escala según un criterio semánticamente vinculado a la naturaleza de la propiedad, de forma tal que hay un estado mayor que todos los otros, uno menor que todos los otros, y los restantes son mayores que al menos uno y menores que al menos uno de los demás.

El tratamiento de las propiedades ordinales tiene mucho en común con el tratamiento de las propiedades categoriales:

1) se adopta un *fundamentum divisionis* adecuado a nuestros intereses cognoscitivos;
2) si los estados son muchos, se debe decidir cuáles se convertirán en categorías de la variable que va a representar la propiedad en la matriz;
3) las categorías se crean y/o delimitan siguiendo los principios de la clasificación;
4) se establecen límites entre las categorías de modo que sean mutuamente exclusivas;
5) se procura que el complejo de las categorías (incluso si es necesario una categoría residual) sea exhaustivo.

Lo que cambia con respecto de las categoriales es que en este caso los estados están ordenados en la realidad (piénsese en el nivel de estudio o en el grado jerárquico), o se puede imaginar un criterio intersubjetivo (es decir, aceptable por cada persona razonable y competente) para ordenarlos.

Obviamente, el mismo orden que percibimos entre los estados debe ser reproducido en los códigos (en este caso, necesariamente numéricos) que se atribuyen a las categorías. En términos técnicos, se dice que la serie de los códigos debe tener una relación monotónica directa con el orden de los estados en la realidad: al crecer la posición ordinal del estado, hay que acrecentar también el código numérico asignado a la modalidad correspondiente.

Habitualmente a las categorías elegidas se les asignan como códigos la serie de los "números naturales"[26] (enteros positivos: 1, 2, 3...) con el supuesto de que ésta sea la asignación menos arbitraria. Pero, si pretendemos aplicar a estas variables sólo técnicas de análisis para variables ordinales, cualquier asignación monotónica directa[27] es equivalente.

Si en cambio se intenta aplicar a estas variables ordinales las más poderosas técnicas para variables cardinales, la distancia entre códigos se convierte en relevante; en matemática, la distancia entre 1 y 2 no es la misma que entre 1 y 5. En este caso, asignar los números naturales como códigos a las categorías equivale a afirmar que las imaginamos todas más o menos equidistantes. Es posible que lo sean, pero no está garantizado.

Como observa el metodólogo norteamericano Edward Tufte, si nuestro conocimiento de la naturaleza de la variable nos dice que las categorías no son equidistantes, asignar los números naturales porque esto es "objetivo" o "no arbitrario", por el contrario introduce un sesgo, y la asignación de códigos "debe incorporar el conocimiento sustancial que el investigador tiene de la propiedad en cuestión" (1970: 440-1). Por ejemplo, si estamos convencidos de que respecto del título de estudio la distancia entre licenciatura y educación secundaria es mayor que la distancia entre educación secundaria y educación primaria, se deben asignar, por ejemplo, códigos como 1 a la educación primaria, 2 a la educación secundaria y 4 a la licenciatura y no la serie de los números naturales como se hizo —para no adelantar el tema— en el ejemplo del apartado 7.1.[28]

La diferencia entre uno y otro criterio de codificación puede tener un impacto considerable sobre coeficientes[29] cardinales si el número de categorías es bajo —como lo es en la mayoría de las variables ordinales. En efecto, las propiedades discretas ordinales que se encuentran en la realidad no tienen un gran número de categorías.[30] Por consiguiente, el problema de unir categorías semánticamente parecidas antes de analizar las relaciones con otras variables, que era central para las variables categoriales, pierde su importancia. Y eso no sólo porque las categorías de una propiedad ordinal habitualmente no son muchas, sino también porque ellas pierden parte de su autonomía semántica, en cuanto el interés semántico pasa de las categorías individualmente consideradas al orden de las categorías. Por lo tanto, en una tabla bivariable que incluye

[26] Los enteros positivos se llaman "números naturales" porque son los primeros que aprenden los niños y los únicos que conocían muchas poblaciones primitivas. Además —como han mostrado interesantes experimentos de los etólogos— la conducta de muchos animales no humanos (no sólo mamíferos, sino también aves) muestra que ellos conocen los primeros números enteros y los usan para contar cuando les sirve para sus propios fines.

[27] La serie 1, 2, 3, 4, 5 como la serie 2, 5, 9, 12, 16 como la serie 5, 15, 37, 42, 48.

[28] La misma opinión se encuentra en Corbetta (2003: 94).

[29] Lo ha demostrado Doreian, asignando varias series diferentes de códigos monotónicos a las mismas categorías ordenadas y correlacionando entre sí las variables resultantes (1972). Sobre el concepto de coeficiente véase el apartado 15.1; sobre coeficientes cardinales véase el apartado 15.6.

[30] Una excepción pueden ser los grados de una organización burocrática muy rígida y jerarquizada.

una o dos variables ordinales se tiende a mirar la tendencia global más que las celdas singulares.[31]

El hecho de que basta con poner una categoría en una serie ordenada con otras para reducir su autonomía semántica puede extrañar a muchos. Para convencerlos, se les invita a considerar el ejemplo siguiente. El ministro de educación de Lubolia (un país subdesarrollado imaginario) declara en la televisión nacional: "Para atestiguar los grandes progresos de la educación en este país, me enorgullece declarar que el 34% de los adultos de Lubolia han alcanzado el nivel de educación secundaria. Este porcentaje es exactamente el mismo del Japón, el país más avanzado en la educación pública".

Lo que el ministro olvida decir es que la distribución de los otros niveles es la que se muestra en la tabla 8.4.

Tabla 8.4: *Contraste de dos distribuciones en variables ordinales*

NIVEL DE ESTUDIO	% EN LUBOLIA	% EN JAPÓN
ningún título	41	5
educación primaria	24	22
educación secundaria	34	34
licenciatura	1	32
doctorado	0	7

Aquel 34% tendría pleno sentido, y podría ser comparado con otros porcentajes, si la modalidad "educación secundaria" tuviera plena autonomía semántica; pero dado que es sólo un escalón de esta escala, pierde autonomía semántica y no puede compararse aisladamente con otras cifras aisladas: se deben comparar distribuciones enteras.

Esta reducción de autonomía semántica es más fuerte para las categorías intermedias en el orden,[32] pero se produce también para las categorías extremas. Para comprobarlo, consideremos las distribuciones de las clases sociales en dos países (tabla 8.5).

Tabla 8.5: *Contraste de otras dos distribuciones en variables ordinales*

CLASES SOCIALES	% EN NORLANDIA	% EN LÉXACO
terratenientes y empresarios	10	10
empleados	35	5
comerciantes	25	5
obreros	20	45
campesinos	10	35

[31] Un ejemplo se verá en el capítulo 15.

[32] La ambigüedad semántica de las categorías intermedias no escapó a los payasos, que proporcionan bromas como la siguiente: un payaso relata a otro la noticia de un grave accidente en la autopista, diciendo: "ningún herido". Tras una corta pausa, con expresión de alivio del oyente, el payaso agrega: "Todos muertos".

Terratenientes y empresarios están presentes en el mismo porcentaje en Norlandia y en Léxaco. Pero no cabe duda de que esta clase goza de una situación mucho más segura y tranquila en Norlandia.[33]

Una importante consecuencia directa de la reducción de autonomía semántica por parte de las categorías ordenadas es que en las tablas que presentan la distribución de los porcentajes relativos a cada modalidad singular deben ser reemplazados, o al menos acompañados, por PORCENTAJES ACUMULADOS. El porcentaje acumulado relativo a una modalidad se calcula teniendo en cuenta no sólo la frecuencia de aquella modalidad sino también las frecuencias de las modalidades que la preceden en el orden. A veces puede ser útil en el análisis presentar PORCENTAJES RETROACUMULADOS, que suman la frecuencia de aquella modalidad con los de las modalidades que siguen en el orden.

La tabla 8.6 reproduce la tabla 8.4 con porcentajes acumulados y retroacumulados, mostrando que con ninguno de ellos el ministro de Lubolia tenía base numérica para jactarse.

Tabla 8.6: *Dos distribuciones con porcentajes acumulados y retroacumulados*

NIVEL DE ESTUDIO	LUBOLIA (% ACUM.)	JAPÓN (% ACUM.)	LUBOLIA (% RETR.)	JAPÓN (% RETR.)
ningún título	41	5	100	100
educación primaria	65	27	59	95
educ. secundaria	99	61	35	73
licenciatura	100	93	1	39
doctorado	100	100	0	7

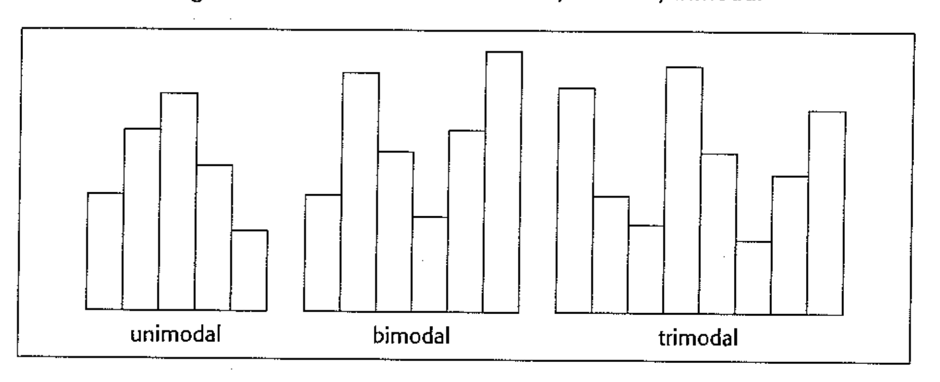
Si los porcentajes fuesen presentados de forma correcta, el ministro estaría forzado a declarar: "En Lubolia el 99% tiene educación secundaria o menos, mientras que en Japón el 61% está en la misma situación" (porcentajes acumulados). Si se basaba en porcentajes retroacumulados, su mensaje habría sido todavía menos contundente: "En Lubolia el 35% alcanzó educación secundaria o más, mientras que en Japón el 73% está en la misma situación".

Una ventaja del recurso a distribuciones con porcentajes acumulados es que esto permite detectar inmediatamente la categoría mediana de una distribución. En una secuencia de cifras, la mediana es la cifra que divide la serie en dos partes iguales, es decir que tiene el mismo número de cifras en ambos lados de la serie (el mismo número de cifras más bajas y de cifras más altas). Cuando en lugar de una serie de cifras se habla de una distribución de frecuencia, la categoría mediana es donde cae el porcentaje acumulado del 50%. Con las distribuciones de la tabla 8.6, ella sería la educación primaria en Lubolia y la educación secundaria en Japón.

[33] Tampoco la reducción de autonomía semántica de las categorías extremas de una escala escapó a los payasos, como lo muestra otra broma clásica de su repertorio. Un payaso anuncia a su compañero otro accidente —esta vez de ferrocarril— agregando que las consecuencias más graves se sufrieron en el último vagón. El otro payaso comenta: "siempre es así: el último vagón es el más perjudicado. ¿Por qué no lo quitan?"

Una segunda consecuencia de la reducción de autonomía semántica de las categorías ordenadas es que lo interesante de la moda (la modalidad con la frecuencia más alta) es su posición en el orden con respecto a las otras modalidades. Además, la presencia de una o más modalidades con frecuencias poco inferiores a la de la moda no reviste interés si ellas se encuentran cerca de la moda (en tal caso la distribución se dice unimodal). Pero si sus posiciones son alejadas, se consideran modas tanto como la más alta, y se habla de distribución bimodal o trimodal. La figura 8.1 muestra un ejemplo de cada tipo de distribución.

Figura 8.1: *Distribuciones unimodal, bimodal, trimodal*



8.3. Propiedades discretas cardinales y variables cardinales naturales

El título de este apartado puede parecer extraño si no se considera el hecho de que todas las variables son discretas, puesto que el continuo (números reales, con infinitas cifras) no se puede representar en la matriz: por lo tanto la expresión 'variable continua', que a menudo se escucha en lugar de 'variable cardinal', es una *contradictio en adiecto* (el adjetivo contradice al sustantivo), así como la expresión 'variable discreta' es redundante. Ya que una variable cardinal también se deriva de una propiedad continua mensurable a través de medición (véase el apartado siguiente), para distinguir las cardinales que se originan a partir de una propiedad discreta se agrega el adjetivo 'naturales', en el sentido de que en estas variables ya sea el punto cero (ausencia de la propiedad) ya sea la unidad de cómputo (el número 1) son naturales y no radican en una decisión del investigador o de la comunidad científica.

Eso no quita que en la fase del análisis, la diferencia entre variables cardinales naturales y variables cardinales métricas (las que derivan de una medición) pierda toda su importancia, y se puede hablar de variables cardinales sin más.

Como se decía al inicio de este capítulo, entre los estados de una propiedad discreta cardinal se pueden legítimamente establecer cocientes: se puede decir que una familia de 8 individuos tiene un número doble de componentes que

una familia de 4.[34] Habitualmente, estas cifras que vienen del mundo real también se usan como códigos numéricos de los estados en el vector relativo. Hay pocas excepciones, dictadas por el sentido común: por ejemplo, el presupuesto de un Estado no se registra en unidades de cálculo (dólares, euros o pesos) sino en múltiplos (millones de dólares, euros o pesos). Sin embargo, siempre hay un cómputo anterior, que es la definición operativa típica de este tipo de propiedades; si se decide registrar en millares o millones de unidades, esto también será parte de la definición operativa. Además, con la salvedad de casos extremos como el presupuesto y parecidos, en el registro de las informaciones cabe privilegiar la sensibilidad y el detalle: las categorías próximas se podrán agrupar más tarde, cuando se presenten o analicen los datos.

Como se decía acerca de variables categoriales, el criterio que rige la recolección de informaciones (máxima sensibilidad) no es para nada el mismo que rige la presentación de los datos y su análisis (eficacia y controlabilidad intelectual). En el caso de las variables categoriales, la razón principal para agrupar categorías se presenta en la fase de análisis, y radica en la necesidad de reducir el número de celdas en las tablas cuando se analicen sus relaciones con otras categoriales y ordinales; pero la posibilidad de reagrupamiento tiene un límite en el hecho de que cada categoría tiene plena autonomía semántica, y por lo tanto sólo se pueden agrupar categorías semánticamente parecidas.

En el caso de las variables cardinales,[35] la razón para agrupar categorías cuando se presentan los datos es otra, y concierne únicamente a esta fase (por el contrario, en la fase de análisis de estas variables el máximo de detalle acrecienta la precisión de los cálculos). Para entender esta razón, piénsese en una distribución de frecuencia de la variable "número de habitantes" cuando la unidad de registro sea el municipio y el ámbito espacial sea un país entero: puesto que difícilmente dos municipios tienen el mismo número de habitantes, en la distribución habría casi tantas filas como la cantidad de municipios, y en cada categoría la frecuencia sería 1; en raras ocasiones 2.

Para evitar este absurdo, se agrupa a las categorías en clases de amplitud variable conforme a los objetivos de la investigación, y teniendo en cuenta el hecho de que en una distribución de frecuencia con más de una veintena de filas se dispersan las informaciones y se pierde eficacia. Obviamente se agrupan las categorías numéricamente cercanas; por ejemplo, las líneas de una distribución de frecuencia del número de habitantes podrían ser: menos de 1.000, de 1.000 a 3.000, de 3.001 a 5.000, de 5.001 a 10.000, de 10.001 a 20.000, y así sucesivamente.

Al agrupar las categorías no hay otro problema que plantearse más que el numérico: la proximidad semántica no es un problema porque las categorías de una variable cardinal (natural como métrica) tienen una autonomía semántica

[34] El hecho de que se puedan establecer cocientes entre ellos no quita que los estados de una propiedad discreta cardinal puedan ser tratados como categorías de una clasificación. Ellas son mutuamente exclusivas (si el número de componentes es 8, no puede ser 7 o 9 o ningún otro número) y la exhaustividad está garantizada por el hecho de que cualquiera sea el estado en una propiedad discreta cardinal, siempre se encuentra un número que lo representa correctamente.

[35] Tanto para las cardinales naturales como para las cardinales métricas.

muy reducida, que va poco más allá del aspecto meramente numérico. Por otro lado, esta escasa autonomía semántica no se puede descuidar, como muestran las reflexiones siguientes. Se habrá notado que la amplitud de las categorías de arriba no es la misma, sino que es creciente. Esto no es por causalidad, en cuanto se basa en el tipo particular de autonomía semántica de las variables cardinales. Como pasa a menudo, el fenómeno no escapó a la sabiduría popular; piénsese en el dicho: "un recién nacido tiene un cumpleaños cada día, un niño lo tiene cada mes, un adulto lo tiene cada año, y un viejo cada diez años".[36]

En términos metodológicos, esto significa que en muchas propiedades discretas cardinales, las categorías más cercanas al cero tienen una mayor autonomía semántica que las otras. El fenómeno es evidente en el número de hijos: (tener) dos hijos no es simplemente el doble que (tener) uno, y la distancia semántica entre tener ninguno y tener uno es mucho mayor que entre tener dos y tener tres. Por el contrario, la distancia semántica entre 8 y 9 hijos no parece nada más que la diferencia numérica entre 8 y 9.

Como consecuencia de todo eso, las clases iniciales de las variables cardinales tienen habitualmente una amplitud reducida: por ejemplo, ningún hijo / un hijo / dos hijos / de tres a cinco / más que cinco.[37]

Ya que las modalidades de las variables cardinales son números en sentido estricto, se las puede someter a todas las operaciones matemáticas: no sólo se puede detectar la mediana —que se puede también detectar en las variables ordinales porque es suficiente que las categorías estén ordenadas— sino también computar el promedio, y muchos otros VALORES CARACTERÍSTICOS que veremos pronto.

Con anterioridad, cabe subrayar que todo lo que se dirá al finalizar este apartado se refiere también a las variables cardinales métricas (véase el apartado siguiente) e incluso a las que se construyen por medio de algunas técnicas particulares, y que llamaremos cuasicardinales (véase apartado 8.5).

De una variable cardinal no interesa sólo el promedio, sino también su VARIABILIDAD, que habitualmente se representa mediante la dispersión de sus datos respecto de este promedio. La variabilidad es el hito de la aproximación estándar en las ciencias sociales que en el apartado 1.4 hemos llamado método de la asociación. Si los estados en una propiedad no varían en una muestra, ésta se llama constante: se puede tomar nota de este hallazgo si no estaba previsto, pero no cabe recoger informaciones sobre los estados individuales, puesto que son todos iguales.

Dado que los códigos de una variable cardinal son números en sentido estricto, no hay dificultad para calcular un valor que sintetice la dispersión de sus datos en una muestra, o en una población entera. El primer paso es transformar cada dato bruto en su desviación del promedio. El dato X_i (símbolo que signifi-

[36] Se podría argumentar que desde un punto de vista semántico, la edad no tiene un rumbo lineal, sino logarítmico.

[37] Como se decía, esto vale también para las variables cardinales métricas. Un ejemplo puede ser: menos de un año, de uno a 3 años, de 3 a 6, de 6 a 10, de 10 a 15, y así sucesivamente. En rigor, eso conlleva que las variables en cuestión dejan de ser plenamente cardinales y se acercan a las de naturaleza ordinal. Habitualmente, este leve alejamiento de la cardinalidad se descuida, y estas variables se tratan como cardinales sin cuestionamientos.

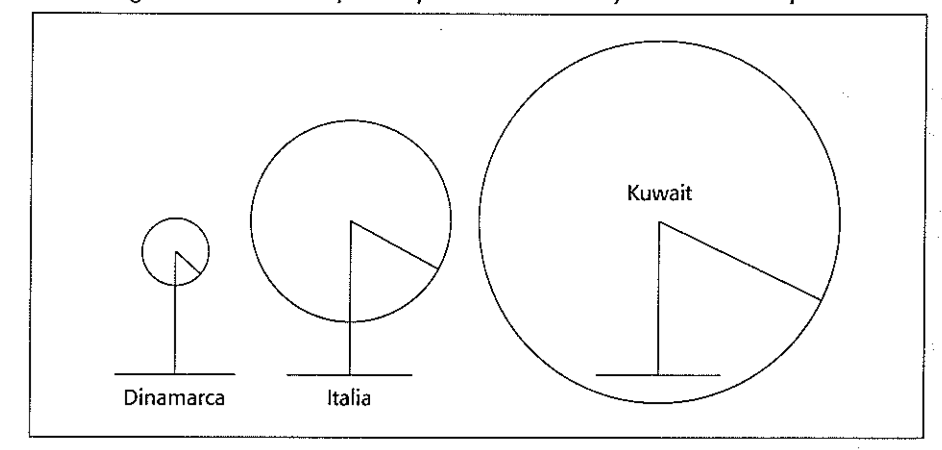
ca un dato cualquiera en la variable X) se convierte en x_i (una manera rápida para escribir $X_i - \bar{X}$),[38] símbolo que significa la desviación de cualquier dato en la variable X respecto del promedio de la misma variable).

Obviamente, unas desviaciones serán positivas, otras negativas. Si las sumamos, se anulan: la suma algebraica de las desviaciones de un promedio es cero por necesidad matemática. Para evitar que la suma se anule tenemos que elevar al cuadrado las desviaciones antes de sumarlas —como sugirió Gauss (véase apartado 7.4). La suma de estas desviaciones cuadradas $\sum x_i^2$ (donde el símbolo $\sum$ significa suma de lo que sigue) en algunos países se llama devianza y es la forma más inmediata de cuantificar la dispersión de los datos de una variable respecto de su promedio. Siendo una suma, su valor depende también del número de sumandos: cualquier distribución tiende a resultar más dispersa cuanto más amplia es la muestra (o la población).

Se sigue que para comparar las dispersiones de dos muestras (o poblaciones) que tienen un número diferente de casos hay que neutralizar el efecto del número de casos —lo que se logra fácilmente, dividendo la devianza por este número.

El valor resultante ($\sum x_i^2/N$, por brevedad s^2) se llama VARIANZA y se puede usar para comparar las dispersiones de la misma variable en dos muestras o poblaciones diferentes.[39] Siendo un promedio de cuadrados, la varianza es una magnitud cuadrática. Para compararla con magnitudes lineales, como el promedio, hay que extraer la raíz cuadrada. El valor resultante ($\sqrt{\sum x_i^2/N}$, por brevedad s) se llama DESVIACIÓN TÍPICA (o desvío estándar), y como veremos se encuentra en muchas fórmulas. La relación entre promedio y desviación típica tiene aplicaciones gráficas inmediatas y fácilmente inteligibles, como la que se muestra en la figura 8.2.

Figura 8.2: *Arbolitos para representar medias y desviaciones típicas*



[38] El símbolo representa el promedio de los datos en la variable X. Nótese la diferencia entre el símbolo del dato bruto (mayúscula) y él de la desviación (cursiva minúscula).

[39] Por ejemplo, si tenemos datos creíbles acerca del ingreso de dos muestras aleatorias y bastante amplias de adultos de dos países, se puede averiguar si estos países tienen niveles diferentes de desigualdad económica.

En esta especie de arbolitos la altura del tronco es proporcional al ingreso medio de tres muestras de ciudadanos (de Dinamarca, Italia y Kuwait). El radio de las copas es proporcional a la desviación típica de los ingresos. Los arbolitos muestran de forma simple y eficaz que los ingresos medios son parecidos, pero en Kuwait hay una gran desigualdad económica, en Dinamarca hay una considerable nivelación, e Italia ocupa una posición intermedia.

Puede también darse la situación exactamente inversa, con desviaciones típicas parecidas y promedios desequilibrados. Por ejemplo si se recolectan datos sobre los sueldos de los empleados de una gran firma y se encuentra que la desviación típica de los sueldos de los cadetes es parecida a la de los sueldos de los gerentes, sería equivocado concluir que el nivel de desigualdad entre los cadetes es el mismo que el que hay entre los gerentes. Esto porque, por ejemplo, doscientos pesos de diferencia en el sueldo mensual son casi nada para un gerente, mientras que son importantes para un cadete. En este caso, los arbolitos tendrían los radios de las copas parecidos, pero las alturas de los troncos muy diferentes.

Para poner remedio a este problema, basta con dividir la desviación típica por el promedio ($s/\bar{X}$), obteniendo el COEFICIENTE DE VARIACIÓN, cuyo símbolo es V. Este coeficiente es una medida absolutamente pura de la dispersión de una distribución, ya que fue depurado de los efectos del número de casos en la distribución y del nivel numérico medio de los datos. Por lo tanto, V es el valor característico más seguro para emplear cada vez que se compare la dispersión de distribuciones diferentes: su nivel de abstracción permite emplearlo también para comparar dispersiones de variables diferentes —obviamente, con tal de que sean variables cardinales.

Los cuatro valores característicos que hemos descrito (devianza, varianza, desviación típica y coeficiente de variación) son valores globales, en el sentido de que se refieren a distribuciones enteras, no a casos singulares. No sirven para confrontar la posición del mismo individuo, o de dos individuos, en las distribuciones de dos variables diferentes. Por ejemplo, si queremos saber si el ingreso de un ingeniero japonés lo pone en una situación de ventaja económica (con respecto del promedio de los empleados japoneses) mayor que el ingreso de un gerente alemán con respecto del promedio de los empleados alemanes.

El procedimiento que permite saberlo se llama ESTANDARIZACIÓN y se realiza en dos pasos: 1) se convierte cada puntaje en su desviación del promedio de los puntajes en la variable, y 2) se divide esta desviación por la desviación típica de la misma variable. El resultado se llama *standard score* (puntaje estándar) y la variable que ha sufrido esta transformación se llama estandarizada.

El puntaje estándar de un caso cualquiera (individuo, agregado territorial, etcétera) en cualquier variable se puede comparar con el puntaje estándar de cualquier otro caso en la misma o cualquier otra variable para saber cuánto —en términos relativos— cada caso se aleja del promedio de la relativa variable. Siguiendo con el ejemplo de arriba, si el puntaje estándar del ingreso del ingeniero japonés es +1,35 y el del gerente alemán es +0,76, eso significa que la posición económica del ingeniero en su país es más ventajosa que la del gerente en el suyo.

La estandarización no sirve sólo para este fin. Sirve también para construir índices aditivos que sumen variables cardinales con una extensión de escala diferente: por ejemplo, para construir un índice de estatus económico, no se pueden sumar directamente el número de celulares que un individuo posee con el número de coches que posee y con su ingreso mensual. Todos estos valores tienen que ser estandarizados antes de sumarlos.

Cabe subrayar que la estandarización, así como todos los valores característicos ilustrados en este apartado, sólo se puede aplicar a variables cardinales (no sólo a las cardinales naturales que derivan de un conteo, a las que se dedicó este apartado, sino también a las cardinales métricas que derivan de una medición, y que se ilustrarán en el apartado siguiente). Pero es tan útil en la construcción de índices aditivos que a menudo se la usa también cuando se trata de crear índices aditivos con variables que no son cardinales, sino cuasi cardinales (un tipo de variables que trataremos en el apartado 8.5). Cuando estas variables son construidas con escalas de extensión diferente, se necesita estandarizarlas antes de sumarlas —y esta necesidad prevalece sobre consideraciones de rigor, que limitarían la estandarización a variables cardinales en sentido estricto.

8.4. Propiedades mensurables y variables cardinales métricas

Como se decía al principio de este capítulo, las propiedades llamadas continuas tienen un número infinito de estados imperceptiblemente diferentes el uno del otro. La mayoría de estas propiedades pertenecen al mundo físico: desde los albores de la civilización, el hombre trató de medir al menos las más elementales (piénsese en las medidas rudimentarias del espacio, del tiempo, del peso, establecidas por muchos pueblos antiguos). Se decía (apartado 1.3) que en la época positivista físicos y químicos hicieron un gran esfuerzo para la universalización de este sistema de medidas y su expansión a muchas propiedades relevantes para sus disciplinas. Un sistema que físicos y químicos de la época presente han sustancialmente heredado, y al cual también se recurre cuando se quiere medir propiedades físicas del hombre, como edad, altura, peso, etcétera.

Otras propiedades (que imaginamos ser) continuas pertenecen al mundo psíquico (piénsese en opiniones, actitudes, valores). Hay que decir "imaginamos ser continuas" dado que no tenemos acceso a los estados en estas propiedades sino a través de manifestaciones exteriores (conductas o declaraciones) del individuo mismo. En las ciencias sociales se observan conductas (especialmente en antropología cultural), y aún más (al menos desde cuando se desarrolló el cuestionario), se solicitan declaraciones. En ambos casos se plantea un problema de inferencia entre lo que se observa y escucha y los estados reales. Esta inferencia es más incierta y compleja si su punto de partida son declaraciones, porque (como se decía en el apartado 7.2) el individuo que declara algo puede tratar —más o menos conscientemente— de presentar una imagen de sí mismo que considera más aceptable, alterando sus estados reales. Sin embargo, un problema de fidelidad (en el sentido que hemos introducido en el capítulo

7) se da también en la observación cuando el sujeto se entera del hecho de que está siendo observado.[40]

Se sigue que estas propiedades psíquicas no se pueden medir en absoluto, en cuanto faltan las condiciones fundamentales para la medición, que veremos a continuación.

La primera condición es que la propiedad sea continua; si no lo es, no se mide: se clasifica (propiedades categoriales y ordinales) o se cuenta (propiedades discretas cardinales). Pero el sociólogo, no teniendo acceso directo a las propiedades psíquicas de los seres humanos, puede sólo imaginar, basándose en la introspección (el análisis de sus propios procesos psíquicos), que algunas opiniones, actitudes y valores varíen de forma continua y no por intervalos discretos. Puede imaginar, pero no puede garantizarlo ni demostrarlo ni siquiera por lo que respecta a sus propiedades; aún menos a las de otros individuos.

La segunda condición es que con anterioridad haya sido establecida y reconocida una unidad de medida. Como ya se dijo, hasta el proceso de universalización de las unidades de medida en las ciencias naturales que se desarrolló en el siglo XIX, sólo existía este reconocimiento para algunas propiedades elementales y en ámbitos espaciales limitados a los Estados, o incluso a provincias; pero la unidad ya no cambiaba en cada acto singular de medición según el capricho de quien medía o según el objeto cuyo estado se medía. Además, a lo largo de los siglos los mercaderes habían establecido tablas de conversión de las medidas locales de longitud o de peso de un lugar al otro,[41] y esto favoreció el proceso de universalización llevado a cabo en el siglo XIX. Una vez acabado este proceso, en las ciencias de la naturaleza se puede hablar de medición sólo cuando se recurra a una unidad de medida universalmente reconocida. La única elección dejada a los que miden es adaptar la unidad a la naturaleza de la propiedad medida, usando oportunamente un múltiplo o submúltiplo de la unidad: por eso, en lugar del metro se pueden usar centímetros, milímetros, micrones, o bien kilómetros, o años luz.

A mediados del siglo XIX los intelectuales interesados por la psiquis humana se planteaban el problema del estatus científico de la naciente disciplina. Algunos psicólogos alemanes (Weber,[42] Fechner, Wundt), que estudiaban en

[40] Después de setenta años se comentan todavía los resultados de los "experimentos" realizados en un establecimiento de Hawthorne (en las cercanías de Chicago) para averiguar cuáles eran los efectos en la productividad de varias condiciones que rodeaban a los empleados en el lugar de trabajo (iluminación, número de escritorios en cada oficina, duración de las recreos, etcétera). Con sorpresa, se constató que los empleados acrecentaban su productividad fuesen cuales fuesen las condiciones, al darse cuenta de que su conducta estaba siendo observada (Roethlisberger y Dikson 1939). A partir de este fracaso del "método experimental" aplicado a seres humanos, los sociólogos empezaron a plantearse el problema del "efecto del experimentador" —un problema que por la gran simpleza de positivistas y conductistas no se había considerado hasta la fecha, suponiendo que los seres humanos reaccionan como objetos físicos.

[41] Tablas de conversión todavía existen entre las medidas anglosajonas de espacio, peso y capacidad y las del resto del mundo.

[42] Naturalmente no estamos hablando del sociólogo Max Weber, sino del psicólogo Ernst Heinrich Weber (1834).

particular el funcionamiento de los cinco sentidos humanos y sus reacciones a estímulos de naturaleza física —por ejemplo las sensaciones de intensidad de luces, sonidos, presiones táctiles, y parecidas— siguiendo el ejemplo de las ciencias físicas trataron de establecer unidades de medida específicas de su dominio, fundadas en "experimentos". Los experimentos consistían, por ejemplo, en someter a los sujetos a sonidos de nivel o intensidad (o a presiones en la piel de intensidad) continua e insensiblemente crecientes. Los sujetos eran instruidos para que manifestasen una señal al experimentador apenas percibiesen una diferencia, un umbral en el proceso de crecimiento del estímulo —proceso que en realidad era continuo.

Por falta de acceso directo a los estados en estas propiedades psíquicas, los psicofísicos[43] tenían que sentar su pretendida unidad de medida en las declaraciones de los sujetos que aceptaban someterse a estas aburridas pruebas de laboratorio. Ya que cada sujeto detectaba y señalaba los varios umbrales en puntos diferentes, el psicofísico que llevaba a cabo aquella particular prueba sobre aquella particular facultad sensorial se encontraba forzado a tomar el promedio de las intensidades efectivas del estímulo en las que cada sujeto había señalado el primer umbral, el segundo umbral, y así sucesivamente.
Tras calcular estos promedios, los psicofísicos bautizaban "diferencia apenas perceptible" (d.a.p.) la diferencia entre los dos, afirmando haber conseguido la deseada unidad de medida.
Por el contrario, se pueden detectar tres importantes diferencias entre la d.a.p. y una unidad de medida en sentido estricto:

a) Dado que la sensibilidad de los seres humanos a estímulos sensoriales de diferente intensidad varía no linealmente sino logarítmicamente,[44] los umbrales medios no se encuentran en posiciones equidistantes a lo largo del *continuum*, sino a distancias crecientes. Por lo tanto hay una d.a.p. entre el primero y el segundo umbral medio, una d.a.p. más ancha entre el segundo y el tercero, y así sucesivamente. Sería como medir el largo de una tabla con una regla de cinco metros, en la cual el primer metro tiene un largo de 80 centímetros, el segundo metro tiene un largo de 86 centímetros, el tercero un largo de 95 centímetros, el cuarto de 107 y el quinto de 120.
b) Como se decía, la serie de estas diferencias apenas perceptibles es establecida sobre la base de declaraciones de una particular muestra de individuos (habitualmente alumnos o conocidos de los psicofísicos). Nada garantiza que cambiando la composición de la muestra (en cuanto a edad, género, etnia, etcétera) la serie de los umbrales medios quedara igual.

43 Psicofísica era y es el nombre de esta especialidad dentro de la psicología.

44 Según la "ley de Fechner" (1860), la sensibilidad tiende a variar con el logaritmo de la intensidad del estímulo (es decir: cuanto más intenso el estímulo, menos fácilmente es percibida su variación).

c) Incluso sometiendo de nuevo la misma muestra a las mismas pruebas, nada garantiza que la serie de los umbrales medios fuese la misma. Aun si se tomase el promedio de los umbrales medios de unas pruebas repetidas, nada garantiza que este promedio sea robusto, es decir estable y definitivo.

Tras valorar el resultado de los intentos de los psicofísicos de medir algunas propiedades psíquicas particulares, cabe destacar que este intento gozaba de condiciones de partida mejores que los sucesivos intentos de otros psicólogos, ya que las propiedades que interesan a los psicofísicos están a medio camino entre la psiquis y la naturaleza física, como el mismo rótulo de la disciplina relativa —psicofísica— nos recuerda. Esto permite cumplir con el requisito de la continuidad al menos de parte del estímulo (sonidos, luces, presiones de intensidad creciente de manera lineal).
En el apartado siguiente se pasará revista a varios intentos de medición de propiedades psíquicas en sentido estricto (opiniones, actitudes, valores) y se comprobará que incluso las técnicas más ingeniosas quedan todavía más alejadas de la medición que las de los psicofísicos. Esto porque no sólo se basan en declaraciones del sujeto, con todas las consecuencias que hemos visto (apartado 7.2), sino también porque —a diferencia de los psicofísicos— los psicólogos, sociólogos, politólogos no tienen manera alguna de someter a los sujetos a estímulos que varíen de forma lineal; es decir, no pueden disponer de una propiedad continua ni siquiera como *input*.

La tercera condición para la medición es la ejecución de un cotejo entre la unidad de medida y el estado en la propiedad que se está midiendo. Cuando se miden propiedades de objetos inanimados o de seres humanos a las que se tiene acceso desde el exterior, este cotejo lo hace el que mide. Pero cuando este acceso directo no se pueda lograr, el investigador puede recorrer tres caminos:

a) observar conductas del sujeto y basándose en éstas inferir su estado; en este caso, no hay unidad de medida ni tampoco cotejo entre ésta y el estado del sujeto en la propiedad;
b) solicitar al sujeto, con preguntas o pruebas de varios tipos, que declare algo que permita al investigador mismo hacer la comparación requerida, con un proceso puramente inferencial;
c) solicitar, con preguntas o pruebas de otro tipo, al sujeto cuyo estado se quiere "medir", que coopere haciendo él mismo la comparación tras proporcionar respuestas que ubiquen su estado directamente en el *continuum* que representa la propiedad.

En el apartado siguiente profundizaremos algunas técnicas que recorren el camino b y el camino c.[45] Por el momento, basta con decir que las diferencias entre

45 De las técnicas basadas en la observación (camino a) se habla en el capítulo 10 de este volumen.

los tres caminos descritos arriba por un lado y un proceso de medición de propiedades físicas por otro lado son tan grandes que no se puede designar a los primeros con el término 'medición' que un uso lingüístico milenario atribuye al segundo.

Descuidando estas diferencias, muchos investigadores (particularmente los norteamericanos y los de inspiración conductista) usan libremente los términos medir/ medida/ medición para cualquier procedimiento de registro de estados (ordenamiento, conteo, construcción de escalas, y a veces incluso para la clasificación). Este verdadero abuso terminológico no tiene otro motivo que el anhelo de legitimación científica mediante imitación ciega de las ciencias físicas; al no poder medir en sentido estricto, se remedia usando de manera impropia todas las palabras que compartan la raíz 'medir'. El mismo fenómeno, y por el mismo anhelo de legitimación, se produjo y todavía se produce con otros términos clave de las ciencias físicas, como 'ley' y 'experimento'.

No es cosa frecuente que los filósofos de las ciencias físicas se ocupen de ciencias sociales; pero cuando encuentran estos abusos terminológicos, no pueden más que comentarlos irónicamente,[46] como la manifestación de un complejo de inferioridad. La misma crítica, con o sin ironía, vino y viene también de parte de los sociólogos y metodólogos de las ciencias sociales más sensibles a consideraciones lógico-filosóficas.[47] Las ciencias sociales lograrán su plena madurez cuando se den cuenta de los límites que la naturaleza de sus objetos pone al empleo de herramientas intelectuales y técnicas, y del hecho de que el estatus científico de una disciplina no depende del empleo impropio de herramientas importadas sino de la conciencia de sus propios límites gnoseológicos y de la capacidad de elegir las herramientas intelectuales y técnicas adecuadas a cada problema.

Para completar este apartado, cabe mencionar algo que no se puede considerar una condición de la medición, pero no puede faltar en cualquier proceso de medición: un criterio para redondear. En efecto, el resultado del cotejo entre un estado en una propiedad continua y la unidad de medida establecida para esta propiedad es un cociente. Ya que es muy improbable que el estado sea exactamente igual a la unidad de medida o a un múltiplo suyo,[48] cuando se compare con la relativa unidad de medida, cada estado se convierte a veces en un número racional (algunas cifras decimales) y casi siempre en un número real (infinitas cifras). Pero no se pueden registrar infinitas cifras en la celda de una matriz, y raramente sirve registrar una o más que una cifra decimal; se sigue que se debe establecer un criterio para redondear.[49]

[46] Véanse, por ejemplo, Cohen y Nagel (1934: 297); Ferguson *et al.* (1938); Carnap (1950: 9); Hempel (1952: 58); Bunge (1973: 105-110).

[47] Véanse, por ejemplo, Torgerson (1958: 14-19); Merton (1940: 17); Kaplan (1964: 213); Berka (1983: 24-34); Duncan (1984: 125-41); Sartori (1984: 31); Roskam (1989: 251); Corbetta (2003: 83-4).

[48] Piénsese en el hecho de que tenemos, por ejemplo, exactamente 35 años sólo durante un minuto, o siquiera una hora; cada otra vez que decimos "35 años" en realidad tenemos algo más o algo menos.

[49] Habitualmente se redondea al número entero más cercano; pero se puede decidir redondear al número entero inferior, y a veces se lo hace.

Eligiendo una unidad de medida y un criterio de redondeo, automáticamente se secciona el supuesto *continuum* en segmentos de amplitud igual a la unidad de medida.[50] Cabe destacarlo porque la división del supuesto *continuum* en segmentos es exactamente lo que tratan de alcanzar las técnicas de construcción de escalas, a las que se dedica el apartado siguiente.

8.5. Propiedades continuas no mensurables y variables ordinales o cuasi cardinales

Como se decía en el apartado 1.4, en el segundo cuarto del siglo XX los psicólogos que se ocupaban de actitudes y opiniones empezaron a plantearse el problema de medir estas propiedades. Inicialmente Louis Leon Thurstone trató de trasladar a este nuevo campo el concepto de diferencia apenas perceptible;[51] después optó por una subdivisión del supuesto *continuum* de la propiedad en diez intervalos juzgados iguales por los sujetos. A pesar de su complejidad, esta técnica no sólo no logra algo parecido al nivel cardinal, sino que no garantiza ni siquiera el nivel ordinal, porque nada en la técnica asegura que los sujetos ordenen de la misma manera las frases que se someten a su valoración. Esto depende del hecho de que las frases planteadas tienen plena autonomía semántica, y por lo tanto cada sujeto interpreta cada una de ellas según sus criterios personales.

Juzgando demasiado complicadas las técnicas elaboradas por Thurstone, en 1932 otro psicólogo norteamericano, Rensis Likert, introdujo algunas drásticas simplificaciones:

1) propuso dividir el *continuum* en un número pequeño de intervalos (inicialmente seis, después sólo cuatro), marcados por siete, y después cinco, frases;
2) renunció a declarar iguales estos intervalos, limitándose a garantizar el orden de las frases que los dividían, tal como los mojones que delimitan un terreno;
3) para garantizar que todos los sujetos compartan este orden, redujo drásticamente la autonomía semántica de lo que se ubicaba en el *continuum* como un mojón: en lugar de plantear frases con pleno sentido pidiendo —como Thurstone— a los sujetos que las ubiquen a lo largo del *continuum,* siguió planteando frases, pero pidiendo a los sujetos reaccionar frente a cada frase con uno de varios grados de (des)aprobación (apruebo plenamente, apruebo parcialmente, estoy indeciso, desapruebo parcialmente, desapruebo totalmente). La técnica ubica en el *continuum* estas cortas expresiones, que tienen una autonomía semántica tan reducida que cualquier sujeto no puede más que compartir su orden (de una plena aprobación a una plena desaprobación).

[50] O a un décimo de ella si registramos una cifra decimal, o bien a un centésimo si registramos dos cifras decimales.

[51] Véase al inicio del apartado precedente.

Las frases a las que el sujeto tiene que reaccionar no le son sometidas aisladamente, sino organizadas en baterías, y comparten el mismo esquema de reacción, de tal forma que el sujeto aprende rápidamente el esquema, lo que permite simplificar su tarea y la del entrevistador, y acelerar todo el procedimiento.

En su versión simplificada, la "ESCALA DE LIKERT", simple y rápida, se difunde en el mundo, convirtiéndose en la técnica de registro de opiniones más conocida y usada en los tres cuartos de siglo que han pasado desde su creación. Por este motivo se le dedicará un examen más profundizado aprovechando décadas de investigaciones en el tema.[52]

La primera observación concierne al hecho de que Likert reintroduce por la ventana la misma cardinalidad que declara haber echado por la puerta. Y lo hace empezando por el título (*Summated Ratings*, valoraciones sumadas) que propone inicialmente para la herramienta que todo el mundo llamará 'escala de Likert'. Las valoraciones que se suman son los números naturales (1, 2, 3, 4, 5) que se atribuyen a los cinco niveles de (des)acuerdo;[53] éstas se suman para asignar a cada sujeto un puntaje total en la escala, que expresa su (supuesto) estado en la propiedad general que la escala pretende medir. Pero no se pueden sumar números ordinales (primero más segundo no es igual a tercero); si se suman, se usan como números cardinales los rótulos numéricos atribuidos a cada una de las simples expresiones de (des)acuerdo.[54]

El empleo de técnicas cardinales con puntajes ordinales no implicaría graves distorsiones si —y sólo si— fuese razonable imaginar que para los sujetos estas cinco expresiones de (des)acuerdo aparecen ubicadas en posiciones aproximadamente equidistantes a lo largo del *continuum*. Pero no es así. Ya algunos metodólogos, basándose en el conocido sesgo en las distribuciones de frecuencia de todas las escalas de Likert,[55] habían apuntado que declararse de acuerdo con cualquier afirmación requiere un empeño psíquico mucho menor que declararse en desacuerdo; esto implica que las posiciones que ocupan en el *continuum* las expresiones del acuerdo y las del desacuerdo no son equidistantes de

[52] Recogidas en Marradi y Gasperoni (2002).

[53] En verdad, en la primera versión de su técnica (1932) había adoptado el llamado "método sigma", ideado por Thurstone, que consistía en atribuir a cada modalidad de respuesta la posición correspondiente a su desviación típica (cuyo símbolo es la letra s o su correspondiente griego, la letra sigma), computada colocando la distribución de frecuencia de las respuestas a una afirmación específica a lo largo de una "curva normal". Pero este criterio fue razonablemente abandonado por Likert en cuanto conllevaba asignar puntajes diferentes a la misma respuesta ("de acuerdo", "parcialmente de acuerdo" y parecidas) según el porcentaje de sujetos que la eligieran, es decir puntajes diferentes en cada investigación.

[54] Más allá del título, Likert usa estos rótulos como números cardinales al sustituir por cinco niveles de (des)acuerdo los siete niveles iniciales; en efecto, él legitima esta sustitución con el alto coeficiente de correlación —una herramienta estadística que sólo se puede usar con variables cardinales— que se encuentra entre las mismas escalas sometidas con cinco y con siete niveles.

[55] El sesgo consiste en el hecho de que, sea cual sea la afirmación planteada, las aprobaciones siempre superan las desaprobaciones (más adelante se verán los motivos). Entre los críticos del supuesto de equidistancia, Jordan (1965) y Galtung (1967).

la supuesta posición central ("indeciso"). Esto ha sido demostrado empíricamente por tres investigadores italianos (Amisano, Rinaldi y Pampanin 2002) aplicando una sofisticada técnica (el análisis de correspondencias) a las respuestas de tres diferentes muestras a muchas baterías de escalas de Likert.[56]

Un segundo problema que presenta la técnica de Likert[57] radica en el hecho de que una afirmación puede ser desaprobada por motivos opuestos. Lo que conlleva que se asignen los mismos puntajes (los que se atribuyen habitualmente al desacuerdo: 0 o 1) a opiniones ubicadas en posiciones muy diferentes a lo largo del *continuum*.

Veamos un ejemplo, sacado de una investigación sobre el nivel de favor de una muestra de italianos para la emancipación femenina. Una de las frases planteadas a los entrevistados era: "Hace falta una ley que asigne a las mujeres el 50% de los empleos". Obviamente se esperaba que los varones machistas, y otros contrarios a la emancipación femenina, desaprobasen. El análisis de los datos comprobó esta expectativa; pero resultó que la oración había sido desaprobada también por algunas mujeres que reaccionando a las otras afirmaciones de la batería se habían mostrado muy favorables a la emancipación. Se recontactó a algunas para pedir explicaciones, y ellas declararon que habían desaprobado la oración porque la mujer debe conquistar la paridad, no esperar que le sea otorgada por ley. Pero siguiendo los criterios normales de codificación, estas campeonas del orgullo femenino recibirían el mismo puntaje que los varones machistas.

Este fenómeno es bien conocido por los que analizan escalas de Likert, y se llama CURVILINEARIDAD,[58] porque asignar la misma puntuación a respuestas que tienen motivos opuestos equivale a doblar el *continuum*, transformándolo en una herradura.

Menos conocido es otro fenómeno que puede distorsionar las reacciones no sólo a las escalas de Likert, sino a todas las técnicas de recolección de opiniones que pidan comentar frases con pleno sentido.[59] El fenómeno emergió escuchando comentarios verbales de entrevistados que respondían frases de escalas de Likert. Habitualmente estos comentarios ocasionales no se graban y se pierden. Por casualidad fueron grabados, y uno de los autores de este libro

[56] En realidad, sus resultados demuestran mucho más: 1) la posición central en el *continuum* no es ocupada por la respuesta "indeciso" sino por las del acuerdo moderado ("parcialmente de acuerdo" y parecidas) ; 2) declararse "indeciso" es percibido por la mayoría de los entrevistados como algo negativo, y por lo tanto se desplaza más cerca a las posiciones del desacuerdo; 3) los entrevistados perciben una distancia grande entre las posiciones del acuerdo (total o parcial), pero poca o ninguna distancia entre las posiciones del desacuerdo.

[57] Como todas las técnicas basadas en frases "no acumulativas" (en inglés *point items*). Sobre el concepto de acumulativo volveremos pronto, hablando de la "escala de Guttman".

[58] Acerca de la curvilinearidad, véanse por ejemplo Edwards y Kenney (1946); Gobo (1997); Pavsic y Pitrone (2003); Palumbo y Garbarino (2004).

[59] Por lo tanto a las "escalas de Guttman" (véase más adelante) y a las técnicas imaginadas por Thurstone (1927; 1928; 1931).

dedicó mucho tiempo a comparar estos comentarios con las reacciones a las mismas frases manifestadas según el esquema previsto por Likert. Con sorpresa descubrió que en ocasiones un entrevistado no reaccionaba al sentido de la oración entera, sino a un solo término —más precisamente, al objeto/sujeto designado por este término. Por este motivo, el fenómeno fue bautizado REACCIÓN AL OBJETO. Una vez descubierto, se le dedicaron investigaciones, que confirmaron su incidencia estimable entre el 5% y el 10% de todas las reacciones a frases de escalas de Likert.[60]

Veamos algunos ejemplos. En una batería dedicada a la visión de la política, una frase era: "Los políticos se interesan sólo por el voto, no por las necesidades de los electores". Un estudiante siciliano de psicología se declaró "totalmente en desacuerdo" y comentó: "Deberían interesarse también por las necesidades, no sólo por el voto. Estoy en desacuerdo con ellos". Por lo tanto, él habría debido aprobar la frase que se le planteaba; pero focalizándose sólo en el objeto de la frase y no en la frase entera, él reaccionó de forma opuesta a la que debía.
En una batería dedicada a la visión de los inmigrantes, una frase era: "Aceptando trabajar por un sueldo reducido, los inmigrantes quitan empleos a los jóvenes italianos". Una docente milanesa jubilada se declaró "en desacuerdo" y comentó: "Estoy en desacuerdo con la política de admitir todos estos extranjeros en nuestro país. Serán la ruina de nuestros jóvenes". Focalizándose en el objeto inmigrantes, ella desaprobó una frase que, dada su opinión en el tema, debía aprobar.
Para detectar cuáles son los motivos principales de esta distorsión, Sapignoli (2002) realizó un "diseño factorial"[61] en el que comparó el efecto de algunas propiedades de los sujetos (edad, género, nivel de instrucción, dimensiones del municipio de residencia) y de algunas propiedades de la frase (número, comprensibilidad y carga emocional[62] de los términos componentes de cada frase). Resultó que la carga emocional del término era la causa de la gran mayoría de "reacciones al objeto": escuchando un término emocionalmente cargado para ellos, algunos entrevistados perdían de vista la oración entera y reaccionaban directamente a aquel término.

Finalmente, el hecho de que las frases sean organizadas en baterías, si bien por un lado acrecienta la rapidez con la que las escalas de Likert se construyen, se someten al entrevistado, son respondidas por él y son analizadas por el investigador, por otro lado reduce la fidelidad de los datos así producidos. Esto se debe a que muchos entrevistados, una vez aprendido el fácil esquema según el cual deben contestar, tienden a considerar todas las frases restantes de la batería como una única afirmación, que aprueban sin más —y asimismo para las frases de las baterías sucesivas. En consecuencia, analizando los cuestionarios que incluyen escalas de Likert, se encuentran a menudo columnas enteras de cruces en las respuestas "de acuerdo" o "parcialmente de acuerdo". Este fenómeno tiene un nombre (*response set* = conjunto de respuestas, dando por descontado que éstas son iguales) y seguramente —considerada la altísima frecuencia con la que se produce— es la falla más grave de las escalas de Likert. Sobre los *response sets* y sus posibles causas se han escrito bibliotecas.[63] En resumen y generalizando, se pueden identificar cuatro motivos:

1) el alto número de frases que componen cada batería,[64] y el hecho de que en un cuestionario a menudo se encuentran muchas baterías, conlleva que en el breve intervalo de una entrevista se someten al entrevistado docenas, tal vez centenares de frases, todas con el mismo esquema de respuestas prefabricadas ("totalmente de acuerdo", etcétera). ¿Cómo sorprenderse si el sujeto termina aburrido por la rutina, y considera todo esto un único trámite, que puede quitarse de encima rápida y fácilmente tras una secuencia mecánica de aprobaciones?
Encontramos aquí la consecuencia del hecho de que los psicólogos —y muchos sociólogos y politólogos junto con ellos— sólo se preocupan por tener muchas variables en sus matrices y poder exhibir altos coeficientes estadísticos de fiabilidad, y descuidan, ya sea la fidelidad de las respuestas, ya sea el respeto que merecen sujetos que no deberían ser aburridos con centenares de preguntas quizá poco relacionadas con sus intereses.
2) Y esto nos lleva a la segunda causa de los *response sets*. Muchos investigadores no se dan cuenta de que a menudo los acervos de conocimientos y las esferas de intereses de los entrevistados no coinciden con los suyos —y si se dan cuenta, no se preocupan, porque lo esencial es rellenar de cualquier forma la matriz de datos. La consecuencia es que muchos entrevistados, no teniendo conocimientos u opiniones sobre los temas que se les presentan, aprueban cualquier oración porque esto es lo que los compromete menos (si desaprobasen, se les podría preguntar por qué).
3) En muchos casos, para impactar al entrevistado y llevarlo a aceptar la entrevista, el entrevistador se presenta como enviado por un instituto académico, a menudo exhibiendo cartas en papel membretado del ins-

[60] Véanse Cacciola y Marradi (1988), Marradi (2002), Sapignoli (2002).
[61] Se llama diseño factorial a un tipo particular de investigación en el que se comparan los efectos de todas las combinaciones de un grupo de variables que se consideran posibles causas de un fenómeno. Véase Corbetta (2003: 145-6).
[62] Para juzgar el nivel de comprensibilidad, Sapignoli usó la fórmula de Flesch (1948), internacionalmente reconocida, adaptada al idioma italiano; para juzgar la carga emocional pidió a un grupo de profesores de psicología y de antropología valorar la carga que cada término en las frases planteadas podía revestir por una muestra de italianos adultos.

[63] Véanse, por ejemplo, Cronbach (1946), Galtung (1967), Schuman y Presser (1981), Converse y Presser (1986).
[64] Este alto número depende del deseo de los psicólogos —en cuya comunidad científica se originaron las escalas de Likert, sucesivamente adoptadas por sociólogos y politólogos— de levantar el valor de los coeficientes estadísticos de fiabilidad (véase el apartado 7.4). La presencia de *response sets* no perjudica el nivel de dichos coeficientes; por el contrario, si las escalas de Likert son suministradas en el marco de un test-retest, este nivel es acrecentado por cada *response set* (véase apartado 7.4).

tituto mismo. Aun informados de que las frases planteadas son opiniones de la gente, la mayoría de los entrevistados —y particularmente los que tienen un bajo nivel de instrucción[65]— las consideran en cambio expresiones de la sabiduría académica y no se atreven a desdecirlas.

4) Como si todo esto no bastara, algunos entrevistadores proporcionan su generosa contribución a la producción de *response sets*. A su descargo se debe anotar que a los entrevistadores que trabajan para una empresa se les paga generalmente poco, se les entrena aún menos, y a menudo deben completar y devolver docenas de cuestionarios en pocos días. No hay que sorprenderse si muchos reducen los tiempos presentando una batería con una oración del tipo: "Ud. está de acuerdo con estas afirmaciones que le voy a leer, ¿verdad?" (Roth 1965; Toller 1994).

Uno de los autores contactó personalmente algunos entrevistadores que habían producido grandes cantidades de *response sets*, y casi todos confesaron cándidamente haber usado siempre fórmulas como la mencionada, agregando que nadie les había enseñado a hacer algo diferente.

El problema de los *response sets* surgió muy tempranamente, y el mismo Likert sugirió la manera aún empleada para evitarlos: mezclar en la misma batería afirmaciones de orientación opuesta, de forma que un entrevistado no pueda declararse razonablemente de acuerdo con todas. Sin embargo, para que este remedio sea eficaz sería necesario que los entrevistadores fuesen adiestrados para actuar apenas el fenómeno se manifieste, llamando con tacto la atención del entrevistado sobre el hecho de que las dos frases que ha aprobado se contradicen y que parece un poco extraño que acuerde con ambas.

Esta intervención sería posibilitada por una serie de conocimientos que no son parte del adiestramiento normal de un entrevistador profesional: la importancia de la fidelidad de los datos; el hecho de que los *response sets* producen datos no fidedignos que son perjudiciales a la encuesta; la función de las frases de orientación opuesta en el detectar los *response sets*; el significado profundo de cada oración que van a plantear, de manera que ellos puedan identificar las frases que se contradicen. Aún más importante que los conocimientos, los entrevistadores deberían haber interiorizado un papel activo y una actitud crítica que no siempre forman parte de los roles y actitudes que se les enseñan en las consultoras.[66]

Si los entrevistadores no actúan para evitar los *response sets* cuando se producen, el remedio de mezclar frases de orientación opuesta termina siendo peor que la enfermedad, porque los entrevistadores registran pasivamente todas las aprobaciones, los codificadores las codifican mecánicamente, y los analistas encuentran muchos datos carentes de sentido que deberían tirar, y que no tiran para no reducir la amplitud de sus muestras.[67] La producción de *response sets* es el lugar donde mejor se manifiesta el juego de suma positiva (desde el punto de vista de los intereses privados) que caracteriza algunos de los sondeos encomendados a firmas comerciales:

[65] Entre otros, Galtung (1967), Schuman y Presser (1981), Converse y Presser (1986), Gasperoni y Giovani (2002) han comprobado que cuanto más bajo es el nivel de instrucción de los entrevistados, más frecuentes son los *response sets* que producen.

[66] Sobre este punto, véanse por ejemplo Roth (1965); Mishler (1986); Bishop *et al.* (1989); Toller (1994); Pavsic y Pitrone (2003).

- los que encargan (sean académicos o empresas) confían en el renombre y en los reaseguros de la empresa de sondeos y ahorran la inversión de tiempo y de atención necesaria para controlar su labor; el único mensaje que se le transmite claramente a la empresa es que se desea recibir los datos lo más temprano posible (véase Boccuzzi 1985, Toller 1994);
- el especialista que diseña el cuestionario lo llena de escalas de Likert porque sabe que es ésta la forma más rápida y simple para recolectar informaciones, codificarlas y cargarlas en la matriz de datos;
- los gerentes de la empresa ahorran el empeño, el tiempo y el dinero necesarios para adiestrar a los entrevistadores, y en particular no les explican que cada oración de las largas baterías debe ser planteada por separado, a fines de evitar que los entrevistadores, dándose cuenta del tiempo necesario para administrar todo el cuestionario, reivindiquen un sueldo adecuado;
- los entrevistadores, sabiendo que la única cosa que interesa a la empresa de sondeos es recibir lo más temprano posible el máximo número de cuestionarios con el máximo grado de (aparente) completud, calculan que para hacer económicamente conveniente su remuneración deben acortar los tiempos de cada entrevista[68] —y la primera víctima de este corte son obviamente las frases individuales de las baterías, que son planteadas en un único bloque, como hemos visto;
- los entrevistados aprenden temprano que aprobar todo lo que se les plantea es la manera mejor para sacarse el trámite de encima, y siguen este criterio a menos que el cuestionario concierna a temas que les interesan particularmente;
- el que analiza los datos, para no reducir drásticamente su número, finge no ver que cada oración de cada batería, cualquiera que sea su texto, resulta aprobada por la gran mayoría de los entrevistados, y que muchos han aprobado todas las frases sin atender a su sentido.

Cada eslabón de esta cadena se adapta a una interpretación reductiva de su papel, que le permite ahorrar empeño y/o tiempo y/o dinero. La única víctima de este juego de suma positiva es la fidelidad de los datos; pero por ésta, como se decía en el apartado 7.4, casi nadie se preocupa.

[67] Véanse las observaciones y los resultados de Gasperoni y Giovani sobre el fracaso del criterio de mezclar oraciones opuestas (2002).

[68] Sobre todo eso atestiguan dos entrevistadoras profesionales (Boccuzzi 1985; Toller 1994). Toller agrega que muchos entrevistadores, para cumplir con los tiempos de entrega requeridos por la empresa, inventan algunas entrevistas, y sugiere desconfiar de los cuestionarios completados de manera demasiado ordenada, porque ellos son casi ciertamente completados por el entrevistador en la calma de su habitación privada.

Un remedio mucho mejor al problema de los *response sets* fue propuesto por Edwards (1954), desarrollando ideas de Rugg y Cantril (1944), y Payne (1949). La idea era bastante simple: en vez de pedir a los entrevistados (des)aprobar frases individuales, pedirles elegir una entre dos frases de orientación opuesta y planteadas juntas, una frente a la otra. Para expresar sus preferencias, los entrevistados son forzados a abandonar su actitud pasiva y a reflexionar sobre lo que significan las dos frases; este aspecto de constricción fue perfectamente percibido por Edwards al llamar *FORCED CHOICE* (elección forzada) a su técnica. Comparando las mismas frases planteadas por un lado singularmente como escalas de Likert y por otro lado en parejas contrapuestas,[69] dos metodólogas italianas han hallado que la tarea de elegir entre dos frases contrapuestas pone de manifiesto la dimensión conceptual apuntada por el investigador, lo que otras técnicas no logran hacer. Proporcionando al entrevistado las herramientas para identificar mejor su posición, se elimina el riesgo de curvilinearidad, se reduce el de reacción al objeto, acrecentando la fidelidad de los datos resultantes (Pavsic y Pitrone 2000: cap. 10).

Es lamentable que una mejora tan importante de la invención de Likert no tenga todavía la difusión que merece en las ciencias sociales,[70] y al momento aparezca muy lejana a reemplazar la versión originaria, como sería oportuno.

Veamos ahora más rápidamente otra técnica, alternativa a la de Likert, que gozó de gran difusión en los años 1950-1970. Su interés radica en el hecho de que su mentor —el norteamericano Louis Guttman— adoptó un criterio radicalmente diferente del de Thurstone y de otros para ubicar frases con pleno sentido a lo largo de un *continuum*: el criterio de la acumulatividad. La idea es simple y se entenderá con algunos ejemplos tomados del deporte y de la escuela. En el salto en alto, si un atleta es capaz de saltar 2,10 metros, se espera que también salte 2,05, 2 metros y todas las alturas inferiores. Puede pasar que falle un salto de una de estas alturas, pero esto es considerado extraño —para Guttman, un "error". La misma cosa con pruebas de matemática: si un alumno soluciona una difícil, el docente espera que sea capaz de solucionar todas las pruebas más fáciles.

Basándose en el criterio de acumulatividad, Guttman imaginó una técnica ingeniosa para ordenar a lo largo de una dimensión conceptual un conjunto de frases con pleno sentido, es decir con plena autonomía semántica. El problema es que si el supuesto de acumulatividad es razonable para pruebas de destreza física y tests de capacidad intelectual, nada garantiza que lo sea también para opiniones. Guttman oculta esta diferencia trasladando sin cuestionamientos a

[69] La misma comparación la hicieron también Bishop *et al.* (1982). Orvik (1972) y Ray (1990) han subrayado que de todas formas no se puede eliminar la deseabilidad social, que lleva al entrevistado a elegir la alternativa que le parece ofrecer la mejor imagen de sí mismo.

[70] Sin embargo, preguntas con elección forzada son empleadas en investigaciones importantes, como por ejemplo los estudios electorales norteamericanos (*National Election Studies*). A mediados del siglo pasado la técnica a elección forzada gozó de un momento de gran difusión gracias a las críticas a la escala F de Adorno *et al.* (1950), de la que se notaba la tendencia a recoger respuestas socialmente deseables y *response sets* (Hyman y Sheatrley 1954; Peabody 1961; Berkowitz y Wolkon 1964).

la esfera de las opiniones los términos relacionados a pruebas (fácil/difícil; pasar/fracasar). Una oración no es fácil o difícil; se aprueba, no se pasa; se desaprueba, no se fracasa. El artilugio terminológico de Guttman no puede esconder el hecho de que su técnica, apropiada para cada género de habilidad, raramente es apropiada para investigar opiniones y actitudes.

Dicho esto, vamos a ver en qué consiste la técnica (conocida como "ESCALA DE GUTTMAN"). Se presentan una serie de frases a una muestra de sujetos, sólo preguntándoles si aprueban o desaprueban cada una de ellas. Luego se organiza una forma de matriz (que Guttman llamaba "escalograma"): en las columnas se ponen las opiniones, ordenadas de la más "fácil" (la que obtuvo más *passes*, es decir aprobaciones)[71] a la más "difícil" (la que obtuvo más *fails*, es decir desaprobaciones). En las filas se ponen los sujetos de la muestra, ordenados de mayor a menor según la cantidad de frases aprobadas. La tabla 8.7 muestra uno de estos escalogramas, que en la intención de Guttman sirven para ordenar[72] ya sea las frases ya sea a los sujetos a lo largo de una dimensión.

> Para facilitar la comprensión, hemos asignado a las frases en la tabla 8.7 las letras del alfabeto en orden de "dificultad" creciente. Por brevedad, cada sujeto de la tabla presenta un patrón particular de (des)aprobaciones. Veamos cómo se le atribuyen las posiciones (y después los puntajes ordinales).
> Ana no presenta problemas: sólo aprobó la oración más fácil, y por lo tanto no cometió errores: su puntaje será 1.
> Consuelo no le gustaría a Guttman, ya que aprobó las frases A, B, C, D y G, desaprobando las dos intermedias. Se debe elegir si asignarle la posición G (y el puntaje 7), considerando "errores" sus desaprobaciones de las frases E y F, o bien la posición D (y el puntaje 4), considerando "error" su aprobación de la oración G. Razonablemente, Guttman sigue el criterio de asignar la posición que minimiza el número de "errores", y por lo tanto a Consuelo se le asignan la posición D y el puntaje 4 y su aprobación de la oración G es considerada error (para señalarlo la letra p se pone en cursiva).

Tabla 8.7: *Un escalograma de Guttman*

FRASES POSICIONES	A 1	B 2	C 3	D 4	E 5	F 6	G 7	H 8	PUNT. ASIGN.	TOTAL *PASSES*	TOTAL ERR.
Ana	p								1	1	0
Consuelo	p	p	p	p			*p*		3	4	1
Diego	p	p	p	*e*	*p*				3 o 5	4	1
Francisco	p	p	*e*	*e*	*p*	*p*			2 o 6	4	2
Josefina	p	p	p	p	*e*	*p*	*e*	*p*	4, 6 u 8	6	2

[71] *Pass*, plural *passes*, es el término inglés para el resultado exitoso de una prueba. Como se decía en el texto, Guttman lo usa en el significado de "acuerdo con una oración".

[72] A diferencia de Thurstone y muchos otros, Guttman se da cuenta de no estar midiendo nada; a diferencia de Likert, él es coherente en subrayar que su técnica sólo puede ordenar oraciones, no medir propiedades subyacentes, rechazando cualquier uso cardinal de los números ordinales que se asignan a las oraciones.

Diego le gustaría todavía menos a Guttman, ya que el criterio de asignar la posición que minimiza el número de "errores" no basta para asignarle una posición y un puntaje. Él aprobó las frases A, B, C y E. Si se considera "error" su aprobación de la oración E, se le asigna la posición C (y el puntaje 3); si se considera "error" su desaprobación de la oración D, se le asigna la posición E (y el puntaje 5).
Una incertidumbre más grave concierne a la posición de Francisco: esa puede ser la B (con puntaje 2) si se consideran "errores" sus aprobaciones de las frases E y F, o bien ser la F (con puntaje 6) si se consideran "errores" sus desaprobaciones de las frases C y D.
Un patrón todavía más complicado presenta Josefina, a la que se pueden asignar tres posiciones diferentes siempre totalizando dos errores: la posición D (y el puntaje 4) considerando "errores" sus aprobaciones de las frases F y H; o bien la posición F (y el puntaje 6) considerando "errores" su desaprobación de la oración E y su aprobación de la oración H; o bien la posición H (y el puntaje 8) considerando "errores" sus desaprobaciones de las frases E y G.

Con la salvedad del primero, todos los patrones mostrados en la tabla presentan "errores". Esto se puso a modo de ejemplo, pero no está alejado de lo que pasa en la realidad. El mismo Guttman (1944), para establecer si un campo semántico es "escalable" o no,[73] introdujo un coeficiente de "reproducibilidad", que consistía sencillamente en la proporción de respuestas "correctas" sobre el total de las respuestas.[74] Si —como a menudo pasaba y pasa— una escala no lograba el umbral deseado en estos coeficientes, Guttman no sacaba la conclusión general de que el supuesto de acumulatividad no es razonable cuando se trata de opiniones, ni la conclusión específica de que no era razonable con respecto a aquel particular campo semántico. Por el contrario, él propuso dos sugerencias, alternativas o paralelas: a) sacar de la muestra a los sujetos culpables de demasiados "errores", ya que "evidentemente" habían contestado sin tener cuidado; b) eliminar de la escala las frases que causaban el número más alto de errores.

La primera sugerencia revela[75] la actitud del autor: si el mundo no se adapta a mis supuestos, hay que cambiar el mundo y no los supuestos. La segunda es ética y científicamente legítima en general: el investigador tiene el derecho, e incluso el deber, de mejorar sus instrumentos, sólo que en el caso específico tiene graves consecuencias técnicas.

[73] Se puede comentar irónicamente que la manera en que Guttman formula el problema es reveladora de la actitud suya (como de muchos otros científicos) respecto de la empresa cognoscitiva: si la escala no funciona como es deseado, la falla está en el campo semántico (es decir, en el mundo) y no en las oraciones (es decir, en las herramientas usadas por el científico).

[74] Como se le observó a Guttman (por ejemplo, Goodenough 1944; Edwards 1957), este coeficiente está sesgado en favor de la "escalabilidad" si muchas oraciones de la escala tienen distribuciones desequilibradas (es decir, con una gran mayoría de *passes* o de *fails*). En este caso, por una razón que se verá pronto en el texto, casi todas las respuestas tienen que ser "correctas".

[75] Una vez más: véase la nota precedente.

En efecto, por una necesidad lógica que veremos pronto, cuanto más central en el *continuum* es la posición de una frase, más probable es que produzca "errores" en una escala de Guttman. Las frases extremas en ambos lados del *continuum* pueden provocar pocos porque en estos casos sólo quien disiente de la mayoría puede estar en "error". La mayoría necesariamente contesta conforme a las expectativas, y por esto no puede estar en error. Por el contrario, si la oración ocupa una posición central en el *continuum*, la expectativa es que las opiniones de la muestra se dividan aproximadamente por mitades. En consecuencia, la posibilidad de que una respuesta no siga el patrón esperado es máxima. Piénsese en la cuerda de un violín: estando bloqueada en las extremidades, en la cercanía de éstas puede vibrar poco; el punto donde puede vibrar más está en el centro.

El hecho de que las frases que eventualmente se eliminan deban tener una posición central en el *continuum* se revela una falla aún peor si se considera que —según las teorías psicológicas más aceptadas— las opiniones sobre cualquier problema tienen generalmente una distribución campanular, y por lo tanto la mayoría de ellas ocupa una posición central —que quedaría desguarnecida si se siguen las sugerencias de Guttman.

Hemos dedicado espacio a esta técnica por dos motivos: a) ella es la progenitora de otras técnicas más complejas,[76] pero basadas en el principio de acumulatividad; b) si aplicada a propiedades para las que resulta razonable el supuesto de acumulatividad, la escala de Guttman logra el objetivo de garantizar un orden compartido de las frases sin reducir su autonomía semántica. Naturalmente esta garantía sólo vale para los sujetos de la muestra, que han contribuido a formar el escalograma. Por otro lado, se debe notar que la escala de Guttman —a diferencia de todas las otras técnicas presentadas en este apartado— en ocasiones puede ser aplicada también cuando la unidad de análisis no sea un ser humano. Por ejemplo Giampaglia (1990) la aplicó con éxito a algunas características de quince municipios rurales en el sur de Italia, encontrando esta secuencia de instituciones (de la más "fácil" —presente en más municipios— a la más "difícil"): una escuela nocturna / una sucursal de banco / un ambulatorio / una escuela profesional / una escuela secundaria / más de un negocio de alimentos / una biblioteca pública / una asociación socorrista / un instituto técnico.

Como subrayaba Guttman, su escala es una técnica meramente ordinal; lo que conlleva que: a) los únicos valores característicos que se pueden aplicar a

[76] Las más conocidas se deben al politólogo holandés Robert Mokken (1971) y al matemático danés George Rasch (1980). La primera se basa en el análisis de las frecuencias en la "celda error", formada por la aprobación de una oración "difícil" y la desaprobación de una oración "fácil". La segunda transforma en complejos procesos matemáticos la intuición básica de que no se puede esperar de manera determinista que un sujeto apruebe todas las oraciones más "fáciles" que su posición en una dimensión y desapruebe todas las oraciones más "difíciles": sólo se puede calcular la probabilidad de que esto se produzca. Todo esto es justo, pero aplicar los cálculos sofisticados de Rasch a respuestas que casi siempre los sujetos dan de forma apresurada y superficial puede dar la impresión de estar disparando a una mariposa con un cañón.

variables construidas con ella son la moda (es decir, la posición en la escala donde se ubican más sujetos) y la mediana; b) para analizar las relaciones de estas variables con otras sólo se pueden aplicar coeficientes ordinales[77] (véase apartado 15.4).

Lo mismo se puede decir para las variables construidas con la escala de Likert. Por otro lado, hay que enfrentarse al hecho de que muchos analistas usan sin cuestionamientos técnicas cardinales con estas variables, porque con esas técnicas pueden fácilmente efectuar un análisis multivariable. Pero esta práctica va a producir resultados poco confiables porque —como ya se vio— las posiciones a lo largo del *continuum* no son equidistantes, y porque las frecuencias están habitualmente mal distribuidas, en el sentido de que las respuestas de acuerdo moderado superan a menudo el 75%, mientras que en caso de equilibrio su cuota sería el 20%. Ésta es la falta más grave, en cuanto —como se verá en el capítulo 15— los coeficientes bivariados son distorsionados cuando las distribuciones de las variables a las que se aplican no son equilibradas.

La sugerencia que cabe dar a quienes quieran aplicar técnicas cardinales es de examinar previamente las distribuciones de todas las variables a las que las van a aplicar, y sólo aplicarlas a las variables que tengan una distribución bastante equilibrada.[78]

Por el contrario, las técnicas cardinales se pueden aplicar sin grave riesgo de distorsión a las variables construidas con las técnicas que vamos a describir en lo remanente de este apartado. La razón es que con éstas se reduce todavía más que con las escalas de Likert la autonomía semántica de las categorías intermedias, que son meras cifras en un intervalo o casillas a lo largo de un corto segmento. El investigador establece el número de cifras en el intervalo o el número de casillas en el segmento,[79] y después comunica al entrevistado la interpretación que se debe dar sólo a las cifras o casillas de los dos extremos de la dimensión. Basándose en estas indicaciones y dentro de estos límites, el sujeto mismo es invitado a manifestar lo que cree ser su estado en la propiedad, eligiendo la cifra o la casilla que le parece corresponder mejor a dicho estado.

La drástica reducción de la autonomía semántica de las categorías intermedias permite eliminar cualquier duda sobre el hecho de que el orden de esas categorías es compartido por todos los sujetos (nadie puede pensar que 8 es menos que 7), y adoptar un razonable supuesto de que las cifras del intervalo y las casillas a lo largo del segmento sean percibidas como equidistantes. Obviamente, si la equidistancia pudiese ser controlada por un observador externo al sujeto entrevistado, se habría cumplido con los requisitos de la cardinalidad y se podría decir que tenemos herramientas para medir opiniones, actitudes y valores. Ya que este control exterior no se puede realizar, la equidistancia sólo puede ser un supuesto, y tenemos que limitarnos a llamar CUASI CARDINALES las variables construidas con estas técnicas.

[77] Naturalmente, si la otra variable es una categorial sólo se pueden usar tablas bivariables y técnicas basadas en ellas, como las log lineares.

[78] Como implica el texto, este análisis previo debería realizarse sobre todas la variables, cardinales y cuasi cardinales incluidas. Sin embargo, para las variables ordinales construidas con las escalas de Likert el criterio a aplicar debe ser más estricto, ya que con ellas no tenemos ninguna garantía de equidistancia de los puntajes, que en realidad sólo son números ordinales. Por otro lado, en el caso rarísimo de una distribución equilibrada de estas variables, tenemos también un indicio indirecto de que las cortas expresiones ("de acuerdo", etcétera) fueron percibidas como aproximadamente equidistantes por los sujetos de aquella muestra, ya que no se manifestó la habitual convergencia en las expresiones de acuerdo moderado.

[79] Como veremos pronto, este número, así como el tipo de interpretación que se propone para los extremos, varía de una técnica a otra.

A fines de moderar cualquier entusiasmo de tinte positivista, cabe adelantar a la presentación de las técnicas una reflexión acerca de todas las diferencias entre éstas y la medición.

> La primera diferencia tiene una naturaleza estructural: en todas estas técnicas se ponen límites a la escala (sus cifras mínima y máxima, o las dos casillas extremas). En cambio, en la medición hay un límite mínimo (el cero) pero ningún límite máximo.[80] Se sigue que la última cifra, o casilla, tiene una carga semántica que no puede existir en la medición, ya que no hay una última cifra. Se puede decir lo mismo para las cifras, o casillas, a medio camino en la escala, mientras que no hay cifras a medio camino en la medición.
> La carga semántica de las cifras, o casillas, extremas y de la central se refleja en las cifras, o casillas, cercanas (que tienen la carga de estar cerca del mínimo, o del máximo, o del promedio). Por este motivo, en este texto se habló siempre de reducir drásticamente, y no de eliminar, la autonomía semántica de las categorías intermedias. En cambio, con la medición (y el conteo) sólo la carga semántica del mínimo puede reflejarse en otras cifras, porque no hay un máximo, ni un promedio; además, ya que la escala no tiene una extensión establecida, no se sabe cuánto son cercanas al mínimo las cifras pequeñas: esto depende de la naturaleza de la propiedad medida. Una edad de 2 años tiene una carga semántica radicalmente diferente de una altura de 2 metros; pero la cifra es la misma: dos.

Estas diferencias estructural-semánticas pueden parecer demasiado sutiles, y por lo tanto menores. Más evidentes son las diferencias gnoseológicas, a las que ya se ha aludido. Como pasa con todas las propiedades psíquicas del ser humano, el investigador no tiene acceso directo a los estados individuales. Además, con esta familia de técnicas el investigador confía al mismo sujeto estudiado la tarea de ubicar su estado directamente en el *continuum* que se imagina represente la propiedad, y que es simulado por la escala.

El hecho de que sea el mismo sujeto quien valora su estado en cada propiedad que el investigador le presenta, y exprese esta valoración con un instrumento al cual no está acostumbrado, conlleva varias consecuencias gnoseológicas bajo la forma de causas de infidelidad:

[80] Exactamente lo mismo se puede decir del conteo. Todo lo que se dice de la medición en este apartado vale también para el conteo.

a) cada sujeto puede concebir de manera diferente del investigador la propiedad en la que debe valorar su estado; de esta concepción diferente no queda ninguna huella en la cifra o casilla elegida;
b) cada sujeto puede equivocarse de buena fe acerca de su estado (por ejemplo, muchos creen ser progresistas mientras que son conservadores);
c) cada sujeto puede alterar conscientemente su estado a fin de dar una imagen de sí que considera más agradable;[81]
d) cada sujeto puede entender mal el mecanismo de funcionamiento de una escala que no le es para nada familiar, es decir la manera de convertir lo que cree ser su estado en una cifra o una casilla;
e) cada sujeto puede comprender en lo esencial el funcionamiento de una escala pero manifestar la tendencia a usarla en modo sesgado (por ejemplo, atribuyendo todos puntajes altos y todos bajos a los objetos a valorar con el "termómetro" —una técnica de la que se hablará pronto).

Sólo a este último tipo de distorsión se puede poner un remedio, y sólo en algunas situaciones particulares que veremos. Para las otras distorsiones no hay ningún remedio, porque tampoco hay modo de detectarlas; considérese también que estas causas de distorsión no son alternativas y pueden darse de forma conjunta.

Volvemos ahora a presentar tres escalas de esta familia. La más antigua es el DIFERENCIAL SEMÁNTICO (Osgood 1952), todavía habitualmente empleado por los psicólogos. Como se ve en la figura 8.3, la escala está formada por siete casillas a lo largo de un segmento, separadas entre sí por un signo de dos puntos. A los extremos del segmento están dos adjetivos opuestos.

Figura 8.3: *Un diferencial semántico*

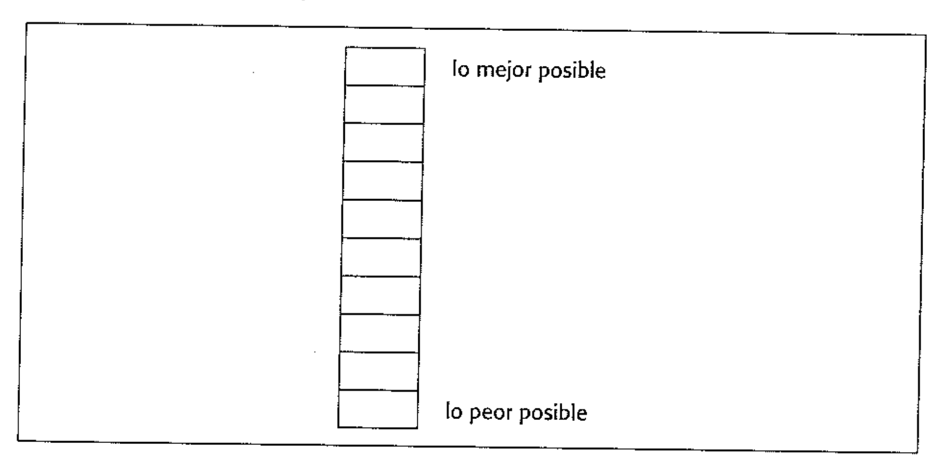
caliente	:___:___:___:___:___:___:___:	frío
honesto	:___:___:___:___:___:___:___:	deshonesto
feo	:___:___:___:___:___:___:___:	bello
duro	:___:___:___:___:___:___:___:	blando
fuerte	:___:___:___:___:___:___:___:	débil
sucio	:___:___:___:___:___:___:___:	limpio
amargo	:___:___:___:___:___:___:___:	dulce
liso	:___:___:___:___:___:___:___:	rugoso

Al sujeto habitualmente se le presentan una docena o una veintena de escalas como la de la figura 8.3, siempre variando las parejas de adjetivos opuestos. Se le pide valorar un referente (puede ser una persona famosa, una institución, un producto, etcétera) eligiendo en cada escala una casilla cuya posición en el segmento manifieste su juicio sobre el referente (por ejemplo, si un político es más "caliente" o más "frío", más "fácil" o más "difícil", etcétera). Si el entrevistado cree que el político X es frío, siguiendo las instrucciones debe poner una cruz en una de las casillas a la derecha del segmento.[82]

[81] Ya se habló varias veces de este fenómeno de "deseabilidad social".

Casi contemporánea al diferencial semántico es otra técnica que sigue el mismo criterio de reducir drásticamente la autonomía semántica de las categorías intermedias. Se presenta al sujeto una columna de diez u once casillas, de las que sólo la más alta y la más baja son interpretadas. En la primera aplicación de esta técnica, Buchanan y Cantril (1953) tomaron una amplia muestra de norteamericanos y pidieron a los entrevistados valorar una serie de etnias o nacionalidades presentes en los Estados Unidos (anglosajones, chinos, italianos, judíos, mejicanos, polacos, etcétera), incluso la del entrevistado, asignándoles una casilla más o menos alta en la escala vertical según su nivel de agrado. En la segunda aplicación, Cantril y Free (1962) bautizaron la casilla más alta de la escala "lo mejor posible" y la casilla más baja "lo peor posible", y pidieron a cada entrevistado ubicar sobre esta escala su situación económica personal y su situación de vida en general (en el presente, cómo era cinco años atrás y cómo pensaba que sería dentro de cinco años) y hacer lo mismo con referencia al país en general (véase la figura 8.4).

Figura 8.4: *Una escala autoanclante*



Ya que era el sujeto mismo quien "anclaba" la escala al establecer la posición de su estado entre los dos extremos, la escala fue bautizada AUTOANCLANTE

[82] Nótese que Osgood recomienda que las parejas de adjetivos sean siempre las mismas, sin tener ninguna relación con la naturaleza del referente que se está valorando (véase Osgood *et al.* 1957). Esto para facilitar la corroboración de una particular teoría suya acerca de la constante presencia de los mismos tres factores en la valoración. Esta teoría es criticable, pero se puede coincidir con Corbetta en juzgar esta técnica como un "aporte al estudio de las dimensiones del espacio semántico" (2003: 253).

—un título que en realidad puede emplearse muy bien como etiqueta para todas las escalas de ese tipo.

Veamos ahora la tercera escala de esta familia, la única importante herramienta de recolección de opiniones que no ha sido producida por psicólogos. Ella nace dentro de un instituto de sondeos académicos que dirigen los politólogos de la universidad de Michigan.[83]

La idea es básicamente una evolución de la escala autoanclante que ya hemos visto, pero aprovecha la imagen de un termómetro, poniendo 100 al máximo y 0 al mínimo de simpatía o aprobación por cualquier objeto que se deba valorar. Por esto ha sido bautizada *feeling thermometer* (TERMÓMETRO DEL SENTIMIENTO), nombre con el cual se ha difundido en el mundo.

Al entrevistado se le presentan en secuencia una serie de referentes (individuos, instituciones, equipos deportivos, etcétera) pidiéndole expresar mediante la escala su grado de simpatía/antipatía. Las instrucciones son elegir y comunicar el puntaje 100 si su simpatía o aprobación es la máxima posible, el puntaje 50 si es neutral, no tiene simpatía ni antipatía por aquel referente, o si su antipatía o desaprobación es la máxima posible. Todos los otros puntajes sirven para expresar un grado de simpatía o de antipatía menor del máximo.

La experiencia ha mostrado que algunos entrevistados no usan estos puntajes intermedios; su conducta presenta una distorsión que puede atribuirse a los tipos (d) o (e) ilustrados arriba: o no comprenden el mecanismo de la escala, o bien lo comprenden pero lo usan de forma sesgada. Un número más alto de entrevistados presenta una distorsión seguramente del tipo (e), en el sentido de que usan también puntajes no extremos pero tienden a dar todos o casi todos puntajes altos,[84] como si aprobasen cualquier cosa. Más raro, pero presente, es el caso de sujetos que tienden a dar todos o casi todos puntajes bajos (como si desaprobasen cualquier cosa) o todos puntajes cerca de 50 (como si no quisiesen comprometerse en una dirección u otra).

Entiéndase: conforme a los objetos que se deben valorar, estas conductas pueden ser perfectamente normales y no sesgadas: por ejemplo, si todos los objetos que se plantean son símbolos de la derecha económica o cultural, es natural que un entrevistado de orientación conservadora los apruebe a todos y uno de orientación progresista los desapruebe a todos. Si un investigador, por negligencia o por cualquier otra razón, somete objetos que son todos símbolos de la misma orientación, él se impide la posibilidad de distinguir a los entrevistados auténticamente conservadores de los que aprueban cualquier cosa, así como la de distinguir a los entrevistados progresistas de los que adrede desaprueban todo.
Por el contrario, se deberían seleccionar atentamente los objetos que se plantean a los entrevistados para valorar, mezclando no sólo símbolos de derecha y de izquierda, sino también objetos percibidos como modernos y como tradicionales, como religiosos y como secularizados, como públicos y como privados, como pertenecientes al centro y a la periferia,[85] de forma que ningún entrevistado pueda razonablemente aprobar o desaprobar todos, y si alguien lo hace, es fácil identificarlo como uno que está usando la herramienta de manera sesgada.

[83] Esta técnica fue empleada por primera vez en el sondeo electoral norteamericano de 1964. En el mismo centro fue desarrollada otra escala de tipo autoanclante. A diferencia de las que hemos visto, esta escala está vinculada a una propiedad específica, la usan los politólogos para registrar la posición que los entrevistados atribuyen a sí mismos o a otros (típicamente personajes públicos) sobre la dimensión derecha-izquierda. La escala está formada por diez u once casillas, y las dos extremas son interpretadas como "derecha" (o "extrema derecha") e "izquierda" (o "extrema izquierda").

[84] Mostrando así la misma propensión a la aquiescencia que hemos visto ser una causa de los *response sets* en las escalas de Likert. El uso sesgado que unos entrevistados hacen del termómetro ha sido notado también por Lau (1982), Wilcox *et al.* (1989).

A condición de que se someta a los entrevistados un número bastante alto de objetos seleccionados según los criterios ilustrados antes, se puede aplicar el remedio presentado unas páginas atrás para las distorsiones del tipo (e). El supuesto es que, si los objetos planteados son muchos y son cuidadosamente variados sobre todas las dimensiones señaladas arriba, el promedio de los puntajes asignados por cada sujeto con el termómetro puede ser considerado un indicio confiable de su tendencia a asignar puntajes más o menos altos independientemente de la naturaleza del objeto valorado.

En otras palabras, se puede descomponer el puntaje asignado en tres partes según la fórmula siguiente:

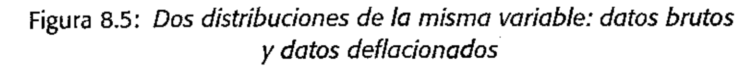
$$p_{xi} = v_{xi} \pm a_{.i} \pm e_{xi}$$

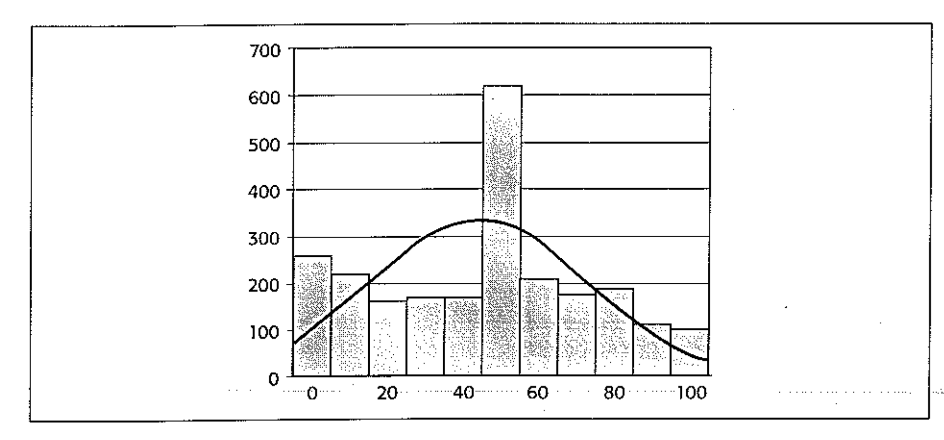
donde p_{xi} significa el puntaje asignado por un sujeto cualquiera (i) a un objeto cualquiera (x);
v_{xi} significa la verdadera actitud del sujeto hacia el objeto;
$a_{.i}$ significa su actitud general a asignar puntajes altos o bajos (el hecho de que se trata de una actitud general está convencionalmente representado en la fórmula por el punto que reemplaza el símbolo **x**);
e_{xi} representa factores erráticos que pueden intervenir en cada proceso de valoración y de reacción a un estímulo.
Estos últimos factores no son controlables; por el contrario, bajo las condiciones subrayadas arriba se puede neutralizar la actitud general (que en la fórmula está representada por $a_{.i}$ y que introduce un sesgo cualquiera en

[85] Algunos ejemplos: en varias pruebas piloto con sujetos italianos, resultó que el clero es percibido como un símbolo religioso, viejo, de periferia, privado, de derecha cultural y en prevalencia de izquierda económica; el Papa es percibido como un símbolo religioso, de centro, público, de derecha cultural; no hay una opinión compartida sobre el hecho de que sea un símbolo nuevo o viejo, y de derecha o de izquierda económico-política. Como siempre pasa, las pruebas piloto brindan sorpresas (y por eso son valiosas para todo investigador que tenga una orientación científica y no cientificista). Baste con decir que en una prueba piloto de los años noventa con una muestra de estudiantes florentinos resultó que la virginidad era percibida como un símbolo nuevo: un probable efecto combinado de la difusión del sida y del renacimiento católico.

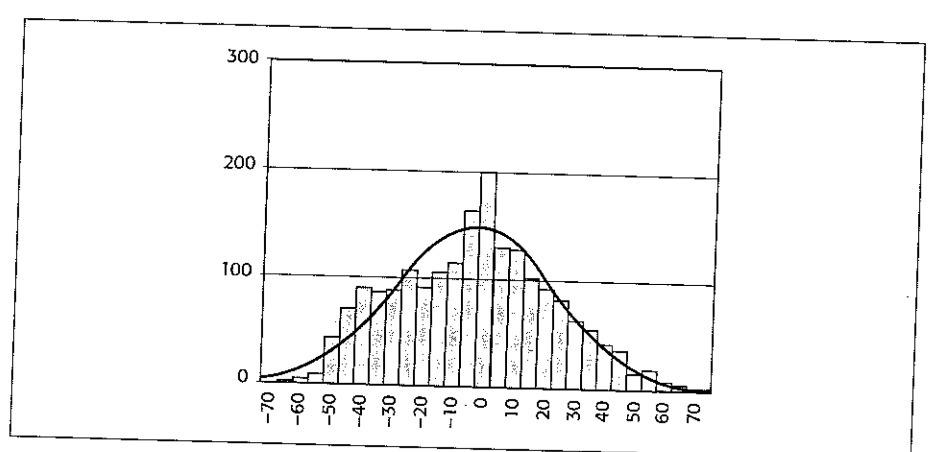
cada puntaje), restando a cada puntaje (p_{xi}) el promedio de todos los puntajes asignados por el sujeto con el termómetro. De esta forma se convierten todos los puntajes brutos en desviaciones de un promedio. Esta conversión de puntajes brutos en desviaciones es lo absolutamente normal en cualquier análisis de datos con variables cardinales; la diferencia radica en el hecho de que todas estas técnicas de análisis computan los promedios por columnas (los puntajes de todos los sujetos de la matriz en una variable), mientras que en este caso los promedios son computados por filas (todos los puntajes que un solo sujeto ha asignado con el termómetro).
Si se quiere también neutralizar la diferente propensión de los sujetos a asignar puntajes extremos o puntajes centrales, se puede dividir cada desviación por la desviación típica[86] de los puntajes asignados. Esta división comprime cada desviación específica en los puntajes de los que prefieren asignar puntajes extremos (muchos 100 o 90 y muchos 0 o 10), y por lo tanto tienen una desviación típica alta; por el contrario, ella ensancha cada desviación específica en los puntajes de los que prefieren asignar puntajes centrales (muchos 50 o 40 o 60), y por lo tanto tienen una desviación típica baja. Por consiguiente, se introduce un mayor equilibrio entre las distribuciones.

Uno de los autores de este libro ideó este procedimiento y lo bautizó DEFLACIÓN por analogía con el procedimiento que usan los economistas para neutralizar el efecto espurio del nivel general de los precios sobre muchas propiedades económicas que les interesan. En su forma completa (cálculo de las desviaciones del promedio y sucesiva división por la desviación típica) la deflación es perfectamente análoga a la estandarización (véase apartado 8.3), con la única diferencia de que se realiza por separado sobre los datos de cada fila, mientras que la estandarización se realiza sobre todos los datos de una sola columna.


Figura 8.5: *Dos distribuciones de la misma variable: datos brutos y datos deflacionados*



[86] Sobre la desviación típica, véase el apartado 8.3.



Fuente: Di Franco (2001: 85).

La figura 8.5 muestra el efecto de la deflación sobre la distribución de los puntajes asignados con el termómetro al partido demócrata cristiano por una muestra de italianos en 1972. Los puntajes brutos se distribuyen de forma lejana de la campanular (la forma que psicólogos y estadísticos esperan para las actitudes y otras propiedades psíquicas); por el contrario, una vez deflacionados, los puntajes se disponen en la clásica forma campanular (muy cercana a la curva normal que la computadora sobrepone a la distribución empírica): pocos individuos con estados ubicados en los dos extremos del supuesto *continuum*, muchos con estados ubicados en la parte central).

8.6. Variables cardinales derivadas

Estas variables forman una categoría aparte, porque no se obtienen con un registro directo, sino con una división entre dos variables cardinales directamente registradas. El cociente entre dos variables cardinales (que pueden ser ambas naturales, ambas métricas, o bien una métrica y una natural) produce otra variable cardinal que no se puede definir ni natural (porque no radica en un conteo) ni métrica (porque no radica en una medición). La definición operativa establece cuál es la variable que se pone como numerador, cuál se pone como denominador, y cuál es el criterio de redondeo (porque un cociente puede producir un número racional con muchos decimales y también un número real con infinitas cifras).

En física se distingue entre variables fundamentales, que se miden directamente, y variables derivadas, que son funciones matemáticas de las fundamentales (por ejemplo, la velocidad es un cociente entre espacio recorrido y tiempo empleado; la densidad es un cociente entre masa y volumen).[87] Por una vez, se

[87] Véanse Campbell (1928), Corbetta (2003: 96).

puede adoptar un término de las ciencias físicas sin temor a distorsionarlo, aplicándolo por analogía a referentes que no le son propios.

Muchos textos de metodología dedican poco espacio a estas variables; sin embargo, ellas desempeñan un papel crucial en las investigaciones cuando la unidad de análisis es un agregado territorial. Lo mostraremos con algunos ejemplos, siguiendo el criterio de la naturaleza de las variables que se ubican como numerador y como denominador respectivamente. Daremos por sentado el hecho de que todas las variables mencionadas han sido registradas en un ámbito temporal adecuado.

Son ejemplos de variables derivadas de un cociente entre dos cardinales naturales:

1) el porcentaje de votos por un partido sobre el total de votos expresados, o de votos válidos;
2) el porcentaje de licenciados sobre el total de individuos con más de 20 años;
3) el porcentaje de juezas sobre el total de los jueces;
4) el porcentaje de médicos sobre el total de la población;
5) el número de desempleados sobre el total de individuos en edad laboral;
6) el número de filiales de bancos por cada mil adultos;
7) el número de camas en hospitales por cada mil residentes.

Son ejemplos de variables derivadas de un cociente entre una cardinal natural y una cardinal métrica:

1) la densidad demográfica (número de residentes por cada kilómetro cuadrado de territorio);
2) el número de coches circulantes por cada kilómetro lineal de rutas;
3) el número de distribuidores de gasolina por cada kilómetro lineal de rutas;
4) el número de tiendas por cada kilómetro lineal de calles o por cada kilómetro cuadrado de territorio urbano;
5) el número de aeropuertos por cada miles de kilómetros cuadrados de territorio.

Son ejemplos de variables derivadas de un cociente entre una cardinal métrica y una cardinal natural:

1) la extensión del espacio habitable por cada residente;
2) la cantidad de energía eléctrica consumida por cada residente;
3) la extensión del espacio verde (parques y jardines) por cada residente;
4) la cantidad de residuos urbanos por cada residente;
5) la cantidad de acero producido por cada empleado en acerías;
6) la cantidad de trigo producido por cada campesino;
7) la extensión de las áreas de estacionamiento por cada coche circulante.

Son ejemplos de variables derivadas de un cociente entre dos cardinales métricas:

1) la extensión del territorio cultivable sobre el territorio total;
2) la extensión de las praderas sobre el territorio total;
3) la extensión total de los parques naturales y áreas protegidas sobre el territorio total;
4) la longitud total de la red de rutas o de ferrocarriles por kilómetro cuadrado de territorio;
5) la longitud total de las costas (en kilómetros lineales) por kilómetro cuadrado de territorio.

Al poder registrar estados sobre propiedades continuas no directamente mensurables, estas variables derivadas tienen una característica típica de los indicadores (el instrumento conceptual al cual dedicaremos el capítulo siguiente). En efecto, la decisión de considerarlas variables correspondientes a propiedades que han sugerido directamente las definiciones operativas apropiadas, o bien indicadores de propiedades más generales, depende del juicio de cada lector. Para posibilitar este juicio vamos a esclarecer el concepto de indicador.

Capítulo 9

INDICADORES, VALIDEZ, CONSTRUCCIÓN DE ÍNDICES

9.1. ¿Por qué indicadores?

En el capítulo 7 se dio por sentado que dar una definición operativa[1] sólo presenta problemas prácticos: encontrar la tabla del censo con los datos requeridos, calcular porcentajes, formular el texto de una pregunta y los textos de las respuestas predefinidas que se ofrecen a la elección del entrevistado, etcétera.

Empero no es siempre así. A veces la naturaleza de la propiedad bajo examen no permite imaginar una definición operativa directa;[2] otras veces ésta se puede imaginar, pero nuestro sentido común, nuestros conocimientos de la situación, nuestra experiencia de investigaciones previas, nos dicen que esa definición operativa sería inadecuada para registrar la propiedad en cuestión.

Imaginemos por ejemplo que nuestra unidad de análisis sea el Estado, nuestro ámbito sea América y la propiedad que interesa sea el nivel de libertad política de que gozan los ciudadanos de cada Estado. Una definición operativa directa sería algo como:

1) busque al señor Estados Unidos, a la señora Argentina, al señor Brasil, etcétera, y pregúnteles cuál es el nivel de libertad política de sus ciudadanos;
2) pídales contestar eligiendo uno de los siguientes niveles: ninguno / escaso / insuficiente / suficiente / bueno / óptimo.

Sin embargo una cosa de este tipo es manifiestamente imposible, porque el señor Estados Unidos, la señora Argentina, el señor Brasil, etcétera, no existen. Se podría imaginar hacer la pregunta al presidente de cada Estado; pero su respuesta no sería confiable, porque incluso el peor dictador declararía que los ciu-

[1] Es decir, establecer el conjunto de acciones, reglas y convenciones que permiten convertir una propiedad de algunos objetos de un cierto tipo en una variable en la matriz de datos.

[2] Es decir, estrechamente vinculada con la propiedad en cuestión y directamente sugerida por ella, sin posibilidad de dudas.

dadanos de su Estado gozan de una plena libertad política. Se registraría la imagen que los presidentes quieren dar del nivel de libertad política de sus países, que es una cosa distinta del nivel en sí mismo.

Una idea aparentemente mejor sería hacer la pregunta a muestras de ciudadanos de aquellos Estados pidiendo sus opiniones. Por el contrario, esta solución presenta problemas aún más grandes, que se pueden sintetizar con la pregunta: "¿la opinión de quiénes?" A la que seguirían otras preguntas, como: ¿Se saca una muestra de todos los ciudadanos adultos o sólo de los que tienen una idea determinada de lo que se entiende por 'libertad política'? ¿Y cómo juzgar qué ciudadanos son efectivamente competentes? ¿Se saca una muestra de todos los ciudadanos adultos o sólo de los que no tienen ningún miedo a declarar su opinión sobre un problema tan delicado a un ajeno desconocido? ¿Y cómo juzgar qué ciudadanos declaran efectivamente lo que piensan?

Además, considerando que las opiniones variarán de un sujeto a otro, ¿cuál se tomará como estado en la propiedad "nivel de libertad política" de cada Estado investigado? Se podría tomar la opinión modal (es decir, compartida por el mayor número de sujetos) o la opinión mediana, si se propuso a los entrevistados una especie de escala ordinal como en la definición anterior.[3]

Al final, incluso admitiendo poder tomar en cada Estado muestras de ciudadanos efectivamente competentes y sinceros en manifestar sus juicios, lo que se registre será sólo la opinión modal (o mediana) de estos ciudadanos acerca del grado de libertad política del respectivo Estado. No se puede negar que tal opinión esté relacionada con el nivel de libertad política en sí; pero tampoco se puede afirmar que sean la misma cosa. Se busca el nivel de libertad política pero se obtiene —en el mejor caso— la opinión modal de ciudadanos competentes y sinceros. Como todas las opiniones, ésta puede ser equivocada, ilusoria.

Imaginemos ahora que nuestro ámbito sea la Argentina, nuestra unidad de análisis sea el varón adulto, y la propiedad que interesa sea su grado de autoritarismo. En este caso, hay un camino sencillo para una definición operativa directa. Sin esfuerzo, se pueden imaginar preguntas como: "¿Cuál es a su juicio su grado de autoritarismo?" y respuestas como: inexistente / bajo / medio / alto. Pero a veces los caminos sencillos no conducen al lugar debido, y éste sería un caso. En efecto, no se registraría el grado de autoritarismo, sino la propensión de cada varón a declararse autoritario —una propensión rara entre los individuos en general, y quizá rarísima entre los varones efectivamente autoritarios.

Se podrían ofrecer otros ejemplos; pero parece mejor tratar de establecer algunos criterios generales:

1) No se puede imaginar una definición operativa directa cuando:
   a) la unidad de análisis es algo que no se puede interrogar (una institución, un agregado territorial, etcétera) y

[3] Nótese que la opinión modal y la opinión mediana coinciden necesariamente sólo si la distribución de las respuestas tiene una forma campanular y simétrica.

   b) los estados en la propiedad investigada no se pueden registrar directamente (como por ejemplo la superficie, etcétera) o sacar de documentos oficiales (como se puede hacer para el número de habitantes, el tipo de sistema electoral, el número de votos de un partido, etcétera).

2) Una definición operativa directa se puede imaginar, pero no es confiable (en el sentido de que va producir muchos datos no fidedignos) cuando la unidad de análisis es un ser humano y la propiedad investigada:
   a) tiene respuestas socialmente deseables (porque casi todos van a afirmar que son honestos y generosos y negar que son autoritarios, obsesivos, tediosos, etcétera);
   b) es algo familiar al científico pero no al lego (no cabe preguntar al hombre de la calle cuál es su grado de anomia, de particularismo, de introversión);
   c) es algo que la moral dominante en la comunidad del entrevistado considera reprobable (como costumbres sexuales atípicas, consumo de drogas, evasión impositiva, etcétera).

   El hecho de que no se pueda dar una definición operativa directa no conlleva una renuncia a recoger informaciones sobre esa propiedad: para hacerlo, hay que buscar una o más propiedades que:
   a) admitan definiciones operativas directas que sean aceptables;
   b) tengan una fuerte relación semántica con la propiedad que interesa.

Estas propiedades se llaman INDICADORES de la propiedad X, y la relación que se establece entre ellos y la propiedad X se llama "relación de indicación". Sobre ella volveremos pronto.[4] Queremos con anterioridad destacar el hecho de que el recurso a indicadores no es específico de las ciencias sociales. Se usan indicadores también en la vida cotidiana, cada vez que algo manifiesto se considera un indicio, una señal de algo que no es manifiesto pero es interesante. La oración "el chico se puso colorado. Eso es un claro indicador de su culpa" no sólo muestra que el concepto es familiar, sino

[4] Muchos autores (Merton 1948; Lazarsfeld y Barton 1951; Galtung 1967; Przeworski y Teune 1970; McKennell 1973; Nowak 1976; Carmines y Zeller 1979) escriben que los indicadores son necesarios cuando el concepto es general y/o tiene gran importancia teórica. Pero, como observó Blalock, "algunas variables teóricamente importantes pueden ser fácilmente operativizadas: la edad, el sexo, la raza, la confesión religiosa" (1961: 163). A esta corta lista se pueden agregar la nacionalidad, el lugar de residencia, el título de estudio, el estado civil, la condición laboral, el tipo de ocupación y —con unidades territoriales— la superficie, la población, la forma de régimen político, etcétera.

Los autores mencionados al inicio de esta nota cometieron un clásico error lógico: de la observación indiscutible que los indicadores son conceptos más específicos de los conceptos que indican (el motivo de esto se verá en el apartado 9.3), sacaron la conclusión de que todos los conceptos generales necesitan indicadores. Pero esto es claramente un *non sequitur*, es decir una deducción equivocada, porque la premisa no conlleva la conclusión.

también que el término existe en el lenguaje ordinario. Y en la tecnología el término se usa para designar instrumentos que revelan el estado de un proceso (indicadores de nivel, de presión, etcétera).

En las ciencias sociales el concepto —no el término— de indicador aparece a mediados del siglo XIX en los escritos de algunos positivistas de habla francesa (Villermé 1840; Quetelet 1869). A finales del siglo, Durkheim habla de *indices extérieurs* que simbolizan hechos interiores y fenómenos morales (1893; 1896). El positivista italiano Niceforo, emigrado a París, dedica una obra a los "índices numéricos de la civilización y del progreso" (1921), usando el término en el mismo sentido que se da ahora a la oración 'indicadores sociales'.
El primero en usar el término 'indicador' en el sentido que se le da en la aproximación estándar a las ciencias sociales fue el sociólogo norteamericano Stuart Dodd (1942). Pero ha sido el metodólogo Paul Lazarsfeld quien codificó este uso, distinguiendo entre 'indicador' e 'índice' (en cuanto combinación de indicadores).[5]

Cabe destacar tres semejanzas entre el concepto ordinario de indicador y el concepto que se usa en la aproximación estándar en las ciencias sociales:

a) en ambos casos el indicador es algo manifiesto (o registrable) que da informaciones sobre algo que no es manifiesto (o directamente registrable);
b) en ambos casos el fenómeno manifiesto / la propiedad registrable puede interesar en sí o como indicador de algo que no lo es. La asertividad de un joven puede interesar en sí o como indicador de un complejo de inferioridad; la tasa de divorcios puede interesar en sí o como indicador de secularización; el porcentaje de votos a un partido de oposición puede interesar en sí o como indicador de insatisfacción con la política del gobierno;
c) en ambos casos el fenómeno manifiesto / la propiedad registrable puede ser considerada —por diferentes observadores e incluso por el mismo— indicador de dos o más fenómenos/propiedades. El hecho de ponerse colorado puede ser considerado un indicador de sensación de culpa o de naturaleza retraída. La tasa de divorcios puede ser considerada un indicador de secularización o de modificación de las estructuras familiares.

9.2. La naturaleza de los indicadores en las ciencias sociales

Las diferencias entre los indicadores que usamos en la vida cotidiana y los que usamos en la aproximación estándar a la investigación social se derivan del hecho de que en esta última el indicador debe su existencia a la necesidad de re-

[5] A los índices se dedicarán los dos párrafos finales de este capítulo.

gistrar propiedades que no se pueden definir operativamente de forma directa. Recuérdese que la definición operativa es el instrumento que nos permite transformar una propiedad del mundo real en una variable que ocupa un vector-columna de la matriz de los datos (véase apartado 7.1). Por consiguiente, la naturaleza de un indicador en las ciencias sociales está estrechamente vinculada con la de los vectores-columna de la matriz. De eso se sigue que en la aproximación estándar a las ciencias sociales:

a) sólo se pueden concebir como indicadores conceptos que se refieren a propiedades;
b) estas propiedades tienen que presentar estados en todos[6] los ejemplares de la unidad de análisis de la investigación, es decir en todos los casos que ocupan las filas de la matriz.

Un ejemplo mostrará mejor lo que se quiere decir. En una clase italiana de metodología, los alumnos tenían que proponer —a título de ejercicio— indicadores de estatus social. La unidad era el individuo adulto, el ámbito era la Italia contemporánea. Fue propuesto el tipo de estación de esquí preferido por el individuo. Pero eso no se puede aceptar como indicador de estatus social porque la mayoría de la población italiana no sabe esquiar, o no tiene medios para frecuentar estaciones de esquí.

Sólo pueden ser elegidos como indicadores (de otro concepto) conceptos que se refieran a propiedades, que sugieran una definición operativa directa y cumplan con el requisito (b) señalado arriba. Naturalmente, éstas son condiciones necesarias pero no suficientes: la condición esencial es que el investigador[7] perciba una estricta relación semántica entre el indicador y el concepto indicado.

Acerca de la naturaleza de esta relación semántica (relación de indicación) se pueden leer las opiniones más variadas.

Muchos autores[8] hablan de una relación causal, generalmente sin establecer si el concepto indicado es la causa o el efecto del concepto indicador. Como veremos enseguida, hay casos en que los conceptos que fueron y son elegidos como indicadores se pueden considerar razonablemente causas del concepto que estarían indicando, y casos en que se pueden considerar efectos; en algunos casos los indicadores que fueron y son elegidos no se pueden considerar ni causa ni efecto del concepto indicado. Aun en los casos en los que se puede imaginar qué relación exista, establecer su dirección no es siempre tan obvio como parece.

[6] Al menos en principio, y con la posible salvedad de unas excepciones marginales.
[7] Naturalmente, esta percepción es sometida al juicio de la comunidad científica. Si la elección no es obvia —como en efecto lo es en la mayoría de los casos— sería de esperar que el investigador la comente para justificarla.
[8] Por ejemplo, Stevens (1951), Blalock (1961), Sullivan (1974), Smelser (1976).

Algunos ejemplos: la tasa de contaminación del aire o la tasa de criminalidad son algunas causas de la percepción de la calidad de vida en un municipio, y a menudo se eligen como indicadores de dicho concepto. Al contrario, la tasa de apoyo hacia un gobierno militar puede considerarse un efecto de una actitud autoritaria más general del sujeto, y a menudo se elige como indicador de autoritarismo.

Veamos un caso menos obvio. En una investigación sobre municipios de Italia central afloró una altísima correlación negativa entre el ingreso per cápita del municipio y la proporción de casas poseídas por la familia que las habitaba. El signo de esta correlación parecía ir en contra del sentido común (cuanto más rica es una familia, tanto más debería estar en condiciones de comprar la casa donde habita). Sin embargo, la relación negativa era tan alta y generalizada que indujo a reflexionar. Y así se comprendió que esa relación no era aleatoria, sino el efecto de una verdadera cadena causal: 1) en los municipios de la montaña, o de alguna otra forma aislados, las oportunidades de trabajo son limitadas; 2) por ese motivo, los ingresos medios son bajos y 3) los jóvenes tienden a desplazarse a municipios urbanos y céntricos. 4) Esta migración interna reduce la demanda de alojamiento en los municipios aislados y pobres y la acrecienta en los municipios ricos y céntricos; 5) por eso, los precios de los alojamientos tienden a bajar en los primeros y a subir en los segundos. 6) Por consiguiente, incluso muchos habitantes pobres de municipios aislados tienen bastante plata como para comprar la casa donde viven, mientras los habitantes más ricos de los municipios céntricos no consiguen hacer lo mismo. Al haber detectado esta cadena causal, el investigador puede estar más confiado en elegir la tasa de casas que son de propiedad de los que las habitan sobre el total de las casas como indicador de subdesarrollo económico.

Pero a veces no se logra detectar la relación que liga una propiedad fácil de operativizar con otra no directamente operativizable, y se elige la primera como indicador de la segunda basándose en una simple asociación estadística. El concepto de asociación será ilustrado en el capítulo 15; ahora baste con un ejemplo, extraído de un renombrado test psicológico, el *Minnesota Multiphasic Inventory*. En ese test, muchas de las 480 preguntas a las que los sujetos tienen que contestar parecen bizarras: un ejemplo es: "¿Siente usted un prurito en su dedos cuando abre un picaporte?"). Es posible que los psicólogos hayan detectado relaciones estadísticas entre sensaciones tan específicas y extrañas e importantes propiedades psíquicas. Basándose en ellas, han decidido elegir tal u otra sensación como uno de los indicadores de autoritarismo (o de anomia, o de inclinaciones paranoicas, u otro).

Naturalmente, éstos son casos límite, que se pueden justificar sólo si necesitan indicadores de un concepto X y no se encuentran conceptos cuya asociación con el concepto X se pueda justificar semánticamente.

Como se decía, los planteamientos acerca de la naturaleza de la relación de indicación son variados. Lazarsfeld (1958; 1966) y otros que lo siguen[9] hablan de una relación probabilística. Pero esto es un burdo error gnoseológico, ya que una relación probabilística puede darse sólo entre un evento (o clase de eventos) y otro evento (o clase de eventos).

[9] Por ejemplo, Verba (1969); Nowak (1976).

Otros[10] hablan de inferencia; esto también es un error gnoseológico, ya que la inferencia es un posible vínculo entre aserciones, nunca entre conceptos.

Guttman (1950) y otros[11] hablan de "muestreo de un universo de contenido". Pero esto es otro error gnoseológico, ya que el muestreo supone que la población (y por lo tanto la muestra) esté formada por individuos (u otros objetos físicos) delimitados, no por entidades vagas y no delimitadas como los conceptos.

Menos grave parece el error de los que hablan de una relación lógica o analítica.[12] En efecto, una relación lógica puede establecerse entre conceptos (por ejemplo la relación entre un género y sus especies es de este tipo); pero la relación entre un concepto y sus indicadores no tiene la misma naturaleza que la relación género-especie, que es necesaria (todos los gatos son felinos, pero hay felinos que no son gatos). Cada investigador, o grupo de investigadores, la establece según su conocimiento de la población estudiada y de la literatura sobre el tema, sus experiencias de investigaciones previas, y en definitiva según su preferencia. Otros investigadores —tanto como el mismo investigador en una ocasión sucesiva— pueden elegir otros indicadores para el mismo concepto, y a menudo lo hacen.

Esto no significa que la elección sea arbitraria. Obviamente el investigador, o grupo de investigadores, tiene el máximo interés en elegir indicadores que sean válidos[13] —es decir que les permitan recoger fielmente al menos algunos aspectos de la intensión del concepto indicado. Por esto no se atreverá a realizar su elección sin conocer la población que va ser objeto de su investigación y la literatura científica sobre el problema que le interesa.[14] No está nada obligado a repetir las elecciones de los que han estudiado el mismo problema de antemano, y todavía menos si estudia el mismo fenómeno en otro ámbito: se asume que el mismo fenómeno pueda tener significados radicalmente diferentes en ámbitos diferentes.[15] Un ejemplo clásico es la tasa de participación electoral,

[10] Por ejemplo, Galtung (1967); McKennell (1973); Singer (1982).

[11] Por ejemplo, Cronbach y Meehl (1955).

[12] Por ejemplo, Nowak (1976); Sullivan y Feldman (1979).

[13] Al concepto de validez se dedicará el apartado 9.4.

[14] Esto no significa que deba elegir sus indicadores en función de su teoría, como sostienen muchos. Sostener que la teoría que se está controlando empíricamente deba dictar, o inspirar, la elección de los indicadores que sirven para controlarla presenta —además de un difuso aire de circularidad— el obvio riesgo de que se elijan los indicadores que ofrecen la mayor probabilidad de corroborarla.

[15] Esta posibilidad no escapó a muchos autores de comparaciones internacionales (Teune 1968; Verba 1969; Frey 1970; Mokrzycki 1983), que subrayaban cómo indicadores del mismo concepto tienen que ser "equivalentes", no iguales. Sin embargo, esta recomendación no es siempre escuchada, y "muchas instituciones que colaboran con los norteamericanos en investigaciones comparadas se quejan de que los cuestionarios que se les envían desde los Estados Unidos son a menudo acompañados por el pedido de que la traducción sea absolutamente literal" (Brislin, Lonner y Thorndike 1973: 39).

que puede ser considerada un indicador de fortaleza de la vida democrática en los países donde hay elecciones libres, mientras que era un indicador de la capacidad de movilización del partido único en los países del Este europeo; e incluso podría ser considerado un indicador de cumplimento de las normas en los países donde votar es un deber jurídico, como la Argentina.

Otra característica de una investigación que puede cambiar radicalmente el significado de un fenómeno es el nivel de la unidad de análisis cuando ésta sea territorial. Un ejemplo de este cambio lo proporciona el diferente significado del porcentaje de ancianos en la población. Si la unidad de análisis es el municipio, esto es un indicador de marginalidad y de estancamiento económico, porque —como se decía antes— los jóvenes tienden a trasladarse a lugares donde encuentran más oportunidades de trabajo. Si la unidad de análisis es más grande e incluye muchos municipios, los municipios periféricos y los centrales se funden en un mismo caso, y por lo tanto el porcentaje de ancianos deja de estar vinculado a la diferencia entre municipios centrales y periféricos. A nivel de esta unidad más grande el porcentaje de ancianos puede ser —en naciones ricas— un indicador de calidad del clima y más en general de la vida, ya que en muchos países los jubilados tienden a trasladarse —si pueden— a las mejores áreas desde estos puntos de vista.

Más allá de estas consideraciones, si un investigador se aleja de las elecciones de indicadores hechas por los que han estudiado el mismo fenómeno en el mismo ámbito y con (una) unidad de análisis del mismo nivel, él debería justificar su rechazo de los indicadores tradicionalmente elegidos para el concepto que está estudiando y su elección de indicadores diferentes.

9.3. Aspectos indicativos y aspectos extraños: la doble pluralidad de las relaciones entre conceptos e indicadores

Al final del apartado 9.1 se observó que una propiedad directamente registrable con una definición operativa puede ser considerada —por diferentes observadores e incluso por el mismo— un indicador de dos o más propiedades. Algunos[16] ven este fenómeno con desaprobación, como una señal de la insuficiente madurez de las ciencias sociales; otros, basándose en una reseña de investigaciones publicadas, se limitan a documentarlo.[17] Trataremos de profundizar este punto destacando sus causas.

> Cuando dos o más personas conversan entre sí, ellas dan por sentado que cada término tiene el mismo significado para todos. Como lo mostraron los etnometodólogos (Garfinkel 1964; 1967), esta convención es necesaria porque si cada interlocutor pidiera el significado de cada término, se abriría un retorno *ad infinitum* y la vida social se paralizaría. Pero si no se tienen

[16] Por ejemplo, Zetterberg (1954), que lamentaba la extrema variedad de los indicadores elegidos en la investigación sobre grupos.

[17] Por ejemplo, el psicólogo Mc Nemar (1946).

> fines prácticos sino cognoscitivos, y se pide a los interlocutores que definan un término cualquiera, casi siempre se encuentra que las definiciones son sensiblemente diferentes —lo que significa que el mismo término designa conceptos distintos para cada sujeto. Por ejemplo, al escuchar el término 'autoritarismo' algunos pueden pensar exclusivamente (o sobre todo) en fenómenos políticos, otros preferentemente en la actitud de los varones con respecto a sus esposas, otros preferentemente en la actitud de los gerentes de una empresa u oficina con respecto a los empleados. Y profundizando más, se encuentra que para algunos el autoritarismo político significa sobre todo el apoyo a un gobierno militar, para otros el deseo de que haya un solo partido, o sólo dos, para otros la opinión de que la toma de todas las decisiones políticas pertenezca al ejecutivo central en desmedro de las autonomías locales, etcétera.
>
> No se pretende decir que cada sujeto da un único y diferente significado a cada término; al contrario, que cada sujeto puede darle muchos significados (algunos centrales, otros marginales, otros ausentes) y la distribución de aspectos centrales, marginales y ausentes varía de sujeto a sujeto. En filosofía del conocimiento, el conjunto de los aspectos centrales y marginales se llama INTENSIÓN de un concepto; toda la argumentación previa se puede sintetizar diciendo que cada término designa —o, al menos, puede designar— conceptos con intensiones diferentes para cada sujeto que lo pronuncia o lo escucha.[18]

Al esclarecer esto, no debería extrañar el hecho de que el mismo concepto A, directamente operativizable, sea elegido por diferentes autores (e incluso por el mismo autor en situaciones diferentes) como indicador de dos o más conceptos no operativizables distintos. Esto deriva naturalmente del hecho de que la intensión de este concepto A varía de sujeto a sujeto y también varía en el tiempo para el mismo sujeto.

> Tomemos como ejemplo la tasa de abortos registrados en cada provincia de una nación de tradición católica donde el aborto es legal. Habitualmente esto se considera un indicador de secularización, es decir, de autonomía con respecto a los preceptos de la Iglesia. En países como la Argentina, donde el aborto no está legalizado, los abortos registrados sólo son un subregistro del total de abortos practicados. La tasa de abortos registrados podría ser un indicador de las malas condiciones sanitarias en que se practican los abortos clandestinos en sectores populares, ya que sólo se registran aquellos casos que derivan en la hospitalización de la mujer en el sistema público. Y en una nación —como por ejemplo Italia, que ha legalizado el aborto pero reconoce a los médicos de los hospitales públicos el derecho a la objeción de conciencia antes de practicarlos—, un investigador podría

[18] En rigor, se podría decir que cada intensión diferente corresponde a un concepto diferente, lo que conllevaría que, mientras que se puede fácilmente averiguar que dos hablantes usan el mismo término, no se puede averiguar —por falta de acceso directo a sus mentes— si están pensando en el mismo concepto.

elegir la tasa de abortos registrados como un indicador de la propensión a la objeción de conciencia de la clase médica de cada provincia, ya que las mujeres, frente a un rechazo (o probable rechazo) de practicarles un aborto, se desplazarían a otras provincias donde la clase médica no tiene problemas de conciencia frente al aborto.

Cada investigador puede considerar central un aspecto diferente de la intensión del concepto de tasa de aborto, y gracias a este aspecto —que podríamos llamar ASPECTO INDICATIVO— elige la tasa de abortos registrados como indicador de diferentes conceptos. En cuanto a los otros aspectos, que no le sirven para establecer la relación de indicación —y que por lo tanto podríamos llamar ASPECTOS EXTRAÑOS—, o los ignora, o los considera no importantes, o se da cuenta de su presencia e importancia pero necesita aquel indicador y espera que los aspectos extraños no tergiversen demasiado la relación de indicación que quiere establecer. Sea cual fuere el reconocimiento que un investigador tiene de los aspectos extraños del indicador que elige, esos aspectos permanecen y contribuyen a influir, por ejemplo, la manera en que una pregunta dada será entendida por los entrevistados. Por ese motivo, la minimización de los aspectos extraños tiene que ser sin duda el primer criterio a considerar en la elección de indicadores. Y dado que no parece posible encontrar un indicador totalmente exento de aspectos extraños, un investigador debería tenerlos muy en cuenta al interpretar sus datos.

A esta pluralidad "desde abajo" en las relaciones de indicación (en el sentido de que un concepto puede ser razonablemente elegido como indicador de otros varios conceptos), le corresponde una pluralidad "desde arriba", en el sentido de que para cada concepto que interesa, pero que no sugiere directamente una definición operativa, cabe elegir más de un indicador.

Veamos el porqué de esto partiendo de un ejemplo. Al inicio de este capítulo se dijo que la libertad política no se puede definir operativamente de manera directa (también en el sentido más específico de "nivel de libertad política que gozan los ciudadanos de varios Estados"). Para permitir un registro de situaciones empíricas cabe descender hacia conceptos más específicos, que son aspectos de la intensión del concepto de libertad política.
Los filósofos políticos convienen sobre al menos tres aspectos de esta intensión: libertad de opinión, de asociación, de prensa. Cada uno de estos aspectos —como otros,[19] cuya pertenencia a la intensión no goza sin embargo de un acuerdo unánime— indica varias direcciones en las que podemos buscar un indicador, es decir, un concepto que se pueda definir operativamente de manera directa y que esté semánticamente vinculado a la libertad política.
El concepto de libertad de opinión sugiere, por ejemplo, el total de días de cárcel sufridos por la expresión de las propias opiniones durante un año

[19] Como por ejemplo la libertad religiosa, o la libertad de ir al extranjero sin restricciones políticas.

dado, dividido por el número de adultos en la población de cada Estado.[20] La definición operativa indicaría dónde proporcionarse la documentación necesaria, cómo efectuar controles consultando agencias internacionales, etcétera.
El concepto de libertad de asociación sugiere, por ejemplo, el número de asociaciones que han sido disueltas —o cuya constitución ha sido prohibida— por motivos políticos (dividido por el número de asociaciones admitidas y no molestadas) durante un período dado.
El concepto de libertad de prensa sugiere, por ejemplo, el total de ejemplares de diarios y periódicos secuestrados por motivos políticos en un año dado dividido por el total de ejemplares circulantes en ese año.
El concepto de libertad de ir al extranjero sugiere, por ejemplo, el número de pasaportes solicitados y no otorgados dividido por el total de pasaportes solicitados durante un año dado.
La definición operativa de los últimos tres indicadores sería análoga a la de arriba (días de cárcel), cambiando obviamente las fuentes de información y la manera de controlarlas.

Como se vio, para posibilitar cualquier registro de informaciones empíricas semánticamente relevantes para el concepto de libertad política cabe recurrir a conceptos mucho más específicos y acotados, que conduzcan directamente a estas informaciones. Por esta razón se habló de un pasaje "desde arriba" del concepto a sus indicadores, y de un pasaje "desde abajo" de cada indicador elegido al concepto relativo. Y ya que los aspectos de la intensión de un concepto general pueden ser muchos, para no reducir demasiado la generalidad de este concepto, y a final de cuentas cambiarlo, es necesario buscar indicadores para muchos aspectos; lo ideal sería encontrar indicadores para cada aspecto relevante de su intensión.[21]

Esto es lo que se entiende cuando se habla de "pluralidad desde arriba" de la relación de indicación. De una forma rápida, la doble pluralidad puede sintetizarse así: cada concepto que no sugiere directamente una definición operati-

[20] Si no se hiciera esta división, los Estados más poblados tendrían una obvia desventaja. Esta operación, que apunta a colocar un fenómeno en sus dimensiones exactas, es denominada 'normalización' y está ilustrada en detalle más adelante.

[21] Esto es solamente un ideal, porque los aspectos relevantes de la intensión de un concepto no son fácilmente enumerables, y además cualquier intento por establecerlos va a ser inevitablemente controvertido. En el apartado 9.5 se aclararán otras dificultades que se encuentran al buscar indicadores para cada aspecto relevante de la intensión de un concepto. Por otro lado, si se dejan sin indicar muchos aspectos importantes de la intensión de un concepto, sería necesario cambiar su denominación. Lo que raramente se hace, como lamentan McNemar (1946), Coombs (1953), Cook y Selltiz (1964), Lutynski (1978). Una consecuencia del no hacerlo la señalan Cronbach y Meehl: "Si el investigador A, para un concepto que denomina 'agresividad', elige indicadores asociados con ataques abiertos a los demás y el investigador B, por otro concepto que también denomina 'agresividad', elige indicadores asociados con la hostilidad reprimida, los mismos resultados que corroboran la teoría del investigador A falsean la teoría del investigador B, y viceversa" (1955, 291).

va necesita una pluralidad de indicadores,[22] y cada concepto que puede ser directamente operativizado puede ser elegido como indicador de una pluralidad de otros conceptos.

Una combinación particular de pluralidad desde arriba y pluralidad desde abajo se encuentra a menudo en el ANÁLISIS SECUNDARIO. Con esta expresión se designa el hecho de que muchos investigadores, para profundizar un problema cognoscitivo que les interesa, en lugar de recolectar directamente datos analizan aquellos ya registrados por otros investigadores y depositados en archivos apropiados.[23]

Naturalmente, cada vez que se analizan los datos de una investigación ajena, estamos limitados a las propiedades que interesaron al autor originario, o autores, de aquella investigación. Por lo tanto, puede pasar que algunas propiedades que interesan no se encuentran en la matriz de datos.[24] Antes de renunciar a considerarlas en sus modelos,[25] muchos investigadores examinan atentamente toda la matriz de la investigación originaria para ver si encuentran una o más variables cuyo concepto correspondiente pueda ser considerado un indicador de una u otra de las variables que les interesan. De esta forma, datos que fueron recogidos para tener información acerca de una propiedad C, interesante en sí o como indicador de una propiedad A, son explotados para tener informaciones acerca de una propiedad B, ésa tampoco directamente operativizable como la propiedad A.

Esta operación —como pasa con todas las relaciones de indicación que un investigador establece— está, al menos en teoría, sometida al juicio de la comunidad científica en cuanto a su validez —un concepto que examinaremos pronto.

9.4. La validez como juicio acerca de la proximidad semántica entre un concepto y su indicador

En rigor, la VALIDEZ es una propiedad del concepto I en cuanto posible indicador del concepto C en un ámbito espacio-temporal determinado con una unidad de análisis dada. El grado en que el concepto I posee esta propiedad no se puede medir,[26] ni averiguar de ningún modo "objetivo". Este grado sólo puede ser evaluado con criterios semánticos por el investigador, el equipo de investigadores, la comunidad de investigadores de una disciplina.

[22] La necesidad de elegir más de un indicador para cada concepto ha sido destacada también por Lazarsfeld: "Cada indicador tiene un carácter específico y no puede ser considerado exhaustivo del significado de otro concepto" (1958, 107).

[23] Estos archivos de datos se difundieron en las últimas décadas en sociología y ciencia política, según el ejemplo de los archivos de relatos y otras informaciones etnográficas que se remontan a fines del siglo XIX, en la época del evolucionismo.

[24] Sobre esto y otros problemas del análisis secundario, véase Hyman (1972).

[25] El modelo es la herramienta básica del análisis de datos en la aproximación estándar. Véase capítulo 15.

[26] Contrariamente a lo que sostienen los metodólogos de inspiración conductista, que propusieron varios procedimientos para "medir" la validez (véase más adelante en este mismo apartado).

Este juicio se da —o se debería dar— teniendo en cuenta los conocimientos del problema y de la población estudiada, los resultados de investigaciones previas, la literatura científica sobre el tema. Una vez registrados los datos, el juicio debe tener en cuenta la fuerza y la forma de la relación (no sólo los coeficientes de asociación)[27] de aquella variable con otras.

En efecto, los metodólogos de inspiración conductista tienden a desacreditar un juicio basado en criterios meramente semánticos (llamado *content validation*, validación de contenido, o —sarcásticamente— *face validation*, validación a simple vista). Ellos juzgan también poco científica una forma de control de validez que usaban los psicólogos de la primera mitad del siglo XX: la validación "por grupos conocidos" (*known groups*). Para ver si un test dado era un indicador válido, por ejemplo, de anomia, se lo aplicaba a una población conocida (habitualmente, los alumnos de una clase de psicología). El test era juzgado válido si el grupo de alumnos que el profesor juzgaba "anómicos" obtenía en el test un puntaje medio significativamente más alto que el resto de la clase (Thurstone y Chave 1929; Frey 1970).

Todas las otras formas de control de validez se basan en coeficientes de asociación entre el vector que brinda los datos del indicador en cuestión y otros vectores de la matriz. Se vio en el apartado 7.4 que también la fiabilidad de los datos se controla —para la mayoría de los investigadores, de inspiración conductista o no— mediante correlaciones de dos a más vectores. Esta analogía ha conllevado una lamentable confusión entre el concepto de fiabilidad (la correspondencia entre situaciones reales y datos en la matriz) y el concepto de validez (el grado de correspondencia semántica entre dos conceptos): algunos usan sin cuestionamientos un término en lugar del otro y confunden estos dos problemas metodológicos distintos.[28] Otros llegan a calcular relaciones matemáticas entre los coeficientes correspondientes y elaboran técnicas para "enfrentar simultáneamente cuestiones de fiabilidad y de validez" (Siegel y Hodge 1968).[29]

Se puede comentar que, en el fondo, ellos son los más coherentes: su idea de validez no es más que la consecuencia de un enfoque que reduce todos los problemas científicos a relaciones matemáticas entre vectores. Una reducción que tiene una justificación oficial —la ciencia tiene que ser "objetiva", y por cierto no hay nada más "objetivo" que la matemática— y una oculta: el deseo de quitarse de encima tareas tan minuciosas y tediosas como controlar el funcionamiento de las herramientas en el campo, ensuciándose las manos con una realidad siempre mucho más compleja que los modelos matemáticos,[30] y también la tarea intelectualmente exigente de reflexionar sobre las relaciones se-

[27] Sobre los coeficientes de asociación, véase el capítulo 15.

[28] Véanse, por ejemplo, Gulliksen (1936); Parry y Crossly (1950); Schuman (1966); Davies (1977); Carmines y Zeller (1979).

[29] Véanse también Cronbach (1949); Bohrnstedt (1970); Allen (1974). Para Lord y Novick (1968) la relación entre validez y fiabilidad es simplísima: la primera es la raíz cuadrada de la segunda.

[30] Sobre la infinita, y continuamente cambiante, complejidad de la realidad ha tenazmente insistido Max Weber (1904; 1906; 1913). Su lección debería ser meditada por los cultores de modelos matemáticos hipersimplificados.

mánticas que tienen dos conceptos en la mente de los individuos investigados en un ámbito espacio-temporal dado.

Veamos ahora las más difundidas de estas técnicas de validación. En la validación "concomitante" o "simultánea" (*concurrent validation*) se calculan coeficientes de asociación con otros supuestos indicadores[31] del mismo concepto (Carmines y Zeller 1979), en la razonable expectativa de que dos o más indicadores del mismo concepto estén positivamente asociados entre sí. Pero incluso una fuerte asociación positiva no demuestra nada de forma conclusiva,[32] porque nada impide que la asociación derive del hecho de que son todos indicadores de otro concepto. Además, si el indicador en cuestión está asociado con algunos y no con otros supuestos indicadores del mismo concepto, no podemos establecer cuál de los dos grupos está formado por indicadores válidos del concepto que nos interesa.

Relativamente más concluyente es la validación "predictiva", en la que se controla si, sobre la base de los puntajes en un indicador, somos capaces de prever correctamente los resultados de una prueba sucesiva (Scott 1968; Carmines y Zeller 1979). Por ejemplo, el vector que brinda los datos del test que trata de detectar la aptitud para conducir un helicóptero, dirigir una orquesta o bailar ritmos afrocubanos se asocia con el vector que brinda los resultados de la prueba final de un curso para conducir helicópteros, dirigir orquestas o bailar esos ritmos. Si la asociación es alta, el test se considera válido como indicador de las habilidades correspondientes, y puede ser adoptado por los directores de los cursos para seleccionar previamente los candidatos. Naturalmente, esta forma de validación sólo puede ser adoptada en situaciones como las descritas, es decir cuando el concepto a indicar puede tener una definición operativa directa bajo la forma de resultados de la prueba final de un curso.

La más interesante de estas validaciones mediante análisis estadístico de la matriz es la "validación de constructo" (*construct validation*), en la que el supuesto indicador no se asocia con indicadores del mismo concepto sino con variables que —según las expectativas teóricas o de sentido común— deberían tener una relación empírica fuerte (positiva o negativa) con el concepto a indicar (Cronbach y Meehl 1955). El indicador se juzga válido si las relaciones son parecidas a aquellas que se esperaban.

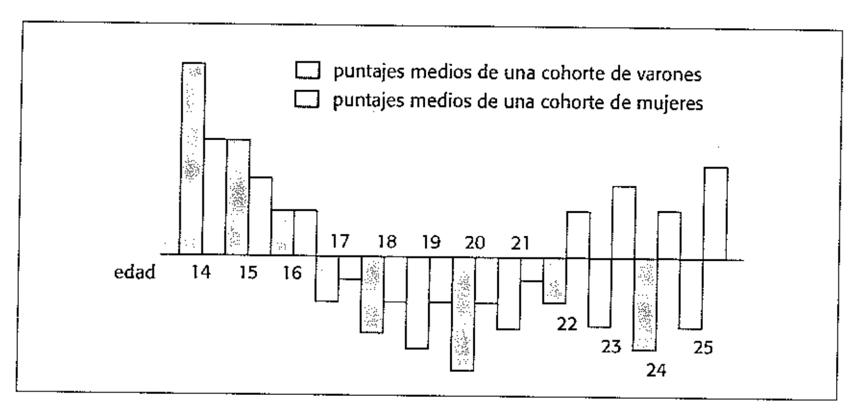
Un ejemplo aclarará lo que se entiende por validación de constructo. En un análisis secundario de una investigación de 7.000 jóvenes italianos en 1971, uno de los autores de este libro efectuó un análisis factorial de una larga batería de escalas de Likert. La primera dimensión extraída no era tan

[31] Como lo observan Frey (1970) y Turner (1979), esto abre el problema de quién validó estos supuestos indicadores. Con otras palabras, la "validación concomitante", y en definitiva todas las formas de validación mediante asociación entre variables en la matriz, abren un regreso *ad infinitum*, al final del cual se encuentra —a pesar de la pretendida "objetividad"— la tan despreciada "validación de contenido": la variable X se considera un indicador válido del concepto C porque está estrechamente asociada con la variable Y; ¿pero quién validó la variable Y en cuanto indicador de C?

[32] Con una feliz expresión debida a Winch y Campbell (1969), una fuerte asociación positiva es *evidence, not proof* (indicio empírico, no prueba).

fácil de interpretar, y fue bautizada "perbenismo" en cuanto distinguía a los jóvenes que aprobaban frases que parecían las recomendaciones de la mamá (juntarse con buenos compañeros, escuchar los consejos de los expertos, evitar posiciones extremas, confiar en los sabios, etcétera) de los jóvenes que las rechazaban.[33] La gran amplitud de la muestra permitió calcular el puntaje medio de cada cohorte de un año en el índice[34] que representaba este factor, dividiendo además entre mujeres y varones. Esos puntajes medios se representan en la figura 9.1. La línea horizontal representa el puntaje 0 (en el análisis factorial, esto es el puntaje medio de la muestra total). Si la columna de una cohorte está arriba de la línea horizontal, el puntaje medio en el índice de aquella cohorte es más alto que el puntaje medio de la muestra; si la columna está debajo de esa línea, el puntaje medio de la cohorte es inferior al promedio de la muestra. Las columnas tienen altura proporcional a la diferencia entre el promedio general y el promedio de aquella cohorte.

Figura 9.1: *Un caso de validación de constructo*



La validación de constructo que aflora claramente de la figura radica en el hecho de que el rumbo de las columnas es el que se esperaría si el índice registrase el verdadero nivel de aceptación que cada cohorte de jóvenes manifiesta para con las recomendaciones de la mamá. Es sabido en efecto que en los países latinos las mujeres se emancipan temprano de la influencia materna, mientras que los varones tardan más en hacerse independientes. El gráfico muestra que este desfasaje se reduce rápidamente, y a los 16 años los varones han alcanzado el mismo nivel de (no)aceptación de las mujeres. A continuación, los niveles de aceptación siguen bajando, pero el de los varones baja más rápidamente: un análisis de regresión múltiple del ín-

[33] El término 'perbenismo' proviene de la expresión *per bene* (como se debe, adecuado a las normas culturales de la "buena burguesía").

[34] En las ciencias sociales se denomina 'cohorte' a un grupo de individuos que tienen la misma edad. No sólo los indicadores, sino también los índices —que son combinaciones de indicadores, y de los que se hablará pronto— son sometidos a un juicio de validez.

dice sobre otras variables mostró que esto depende de la mayor exposición que los jóvenes italianos tenían en aquellos años a influencias exteriores a la familia, como la escuela (que significa profesores, lecturas y —más que otra cosa— compañeros a imitar).[35]
El nivel de aceptación de los varones alcanza su mínimo a los 20 años, y luego sigue siendo muy inferior al promedio general. Al contrario, el puntaje de las mujeres, tras detenerse por debajo del promedio a los 18-20 años, sube nueva y rápidamente por encima del promedio de la muestra, hasta alcanzar a los 25 años un nivel parecido al que tenía a los 14 años. Este rumbo se explica fácilmente si se piensa en el hecho de que las jóvenes italianas de 1971 tendían a casarse mucho más temprano que las de hoy en día, y por lo tanto a los 20 años empezaban a posicionarse ante las frases que hemos mencionado (juntarse con buenos compañeros, escuchar los consejos de los expertos, evitar posiciones extremas, confiar en los sabios, etcétera) no tanto como hijas sino como madres.

Cabe comentar que esta convincente valoración de constructo sólo se pudo lograr analizando en detalle y por separado el puntaje de cada cohorte, cotejando el rumbo de las cohortes femeninas y el de las masculinas. Ésta no se alcanzaría si se correlacionase el índice en cuestión, globalmente considerado, con cualquier otra variable globalmente considerada —como prescriben las técnicas estándar de valoración. El juicio sobre la validez de un indicador o de un índice, aun cuando se base en el análisis de sus relaciones con otras variables en una matriz de datos, a menudo tiene que ir más allá de lo que pueden brindar simples coeficientes globales de asociación.[36]

9.5. Algunos ejemplos de elección de indicadores

Algunos ejemplos ayudarán a comprender mejor la naturaleza de los indicadores y podrán servir como modelos para la elección de nuevos.

Empezaremos con un concepto que se quiere operativizar en prácticamente todas las encuestas sociológicas sobre seres humanos adultos —el de estatus socioeconómico— y que sin embargo no sugiere una aceptable definición operativa directa (no se puede preguntar directamente a una persona de la calle cuál es su estatus socioeconómico porque muchos no entenderían la pregunta, y otros le darían significados muy diferentes de los que se le dan en sociología).

Como ya se dijo, en teoría se deberían buscar indicadores para cada aspecto relevante de la intensión del concepto de estatus socioeconómico. Por otro lado, esta búsqueda encuentra obvios límites en el hecho de que no es siempre fácil encontrar conceptos que sugieran definiciones operativas directas —los únicos idóneos para ser elegidos como indicadores. Además, cuanto más grande es el número de indicadores, más difícil es la construcción de un índice que sintetice —como veremos en el apartado siguiente.

Por estos motivos no se pretende ofrecer una lista exhaustiva de indicadores, pero igualmente se sugiere más de un indicador para el mismo aspecto, o para aspectos cercanos, de forma tal que el lector que quiera inspirarse en la lista tenga un abanico dentro del cual elegir a su vez.

Aspectos relativos a la educación:
- título de estudio del sujeto;
- título de estudio del padre del sujeto;
- título de estudio de la madre del sujeto.[37]

Aspectos relativos a la ocupación:
- ocupación principal del sujeto empleado o última ocupación del sujeto desempleado o jubilado;
- ocupación principal del jefe o jefa de la familia de procedencia.

Aspectos económicos:
- ingreso anual medio del sujeto en los últimos 5 años;
- ingreso anual medio de su familia en los últimos 5 años / número de individuos convivientes;
- metros cuadrados de espacio habitable a disposición de la familia / número de individuos convivientes;
- tipo al cual pertenece la vivienda principal de la familia.

Bienes duraderos poseídos por la familia:
- número de coches poseídos por la familia / número de adultos convivientes;
- número de equipos de TV poseídos por la familia / número de individuos convivientes;
- costo del coche más caro poseído por la familia.

Costumbres del tiempo libre:
- gasto mensual medio (en el último año) para comidas de la familia en restaurantes / número de individuos convivientes;
- gasto medio anual en vacaciones de miembros de la familia / número de individuos convivientes.

[35] Hasta los años 1970 en Italia muchas familias retiraban a las niñas de la escuela a los 14-15 años, y a muchas no se les permitía salir de casa sin parientes o hermanos hasta los 18-20 años. El análisis de regresión múltiple es un desarrollo del análisis de regresión (del que se hablará en el apartado 15.6). Esa técnica permite valorar la influencia relativa de algunas variables en otra, llamada independiente. Todas las técnicas de regresión y correlación sólo se aplican a variables cardinales o cuasi cardinales.

[36] Sobre los coeficientes de asociación, véase el capítulo 15.

[37] El título del padre y el de la madre se pueden combinar en una sola variable, como veremos en el apartado siguiente.

El lector habrá notado que a menudo el indicador incluye la división por el número de individuos o de adultos convivientes en la familia del sujeto. Sólo así el indicador cobra su pleno sentido: se entiende en efecto que poseer tres coches o tres equipos de televisión cambia de significado si el poseedor es solo o si es parte de una familia de ocho miembros. El número de miembros tiende naturalmente a influir los niveles de muchos indicadores cuantitativos, pero no tiene ninguna relación con el concepto de estatus: es un aspecto ajeno a la relación de indicación, y dividiendo por el número de miembros se neutraliza su influencia.

Esta operación de dividir una cifra por una base relevante para neutralizar aspectos de un fenómeno que no interesan se llama NORMALIZACIÓN y se aplica habitualmente no sólo en la creación de indicadores, sino también en el análisis estadístico de los datos.

Pasemos ahora a un concepto que se usa en muchas investigaciones cuya unidad de análisis es la provincia: el nivel de dinamismo económico-demográfico.

Para los *aspectos económicos* en general se podrían elegir indicadores como:

- la variación[38] en el ingreso per cápita;
- la variación en los depósitos bancarios per cápita;
- la variación en el gasto en obras públicas.

Para los *aspectos relativos a empresas* se podrían elegir:

- el saldo demográfico de las empresas (total de nuevas empresas menos total de empresas quebradas) en el intervalo considerado;
- la variación del monto anual de inversiones per cápita;
- la variación del consumo anual per cápita de kW/h para uso industrial;
- la variación en la proporción de suelo ocupado por establecimientos industriales y comerciales.

Para los *aspectos relativos a la ocupación* se podrían elegir:

- la variación en la tasa de empleados sobre la población en edad laborable;
- la variación en la cuota de empleados procedentes de otras provincias o Estados.

Para los *aspectos demográficos* se podrían elegir:

- el saldo demográfico natural (total de nacidos menos total de fallecidos) en el intervalo considerado;
- el saldo demográfico debido a transferencias de otras provincias o Estados menos transferencias a otras provincias o Estados.

[38] Todas las variaciones se entienden de un año X a un año Y, definidos por el investigador según sus intereses.

Considerando que el desarrollo económico se ve favorecido por la difusión de la instrucción, se podría agregar la variación en la cuota de residentes graduados.

Finalmente consideremos un concepto que se puede encontrar en investigaciones con unidades ya sea individuales ya sea territoriales: el grado de emancipación femenina.

Cuando la unidad es territorial (una provincia, un Estado) un abanico de indicadores podría incluir:

- el cociente entre la tasa de ocupación femenina y la tasa de ocupación masculina;
- el total de mujeres con cargos políticos dividido por el total de hombres con cargos semejantes:
- el total de conductoras de autobús dividido por el total de conductores;
- el total de niños menores de cinco años dividido por el total de mujeres de 20-50 años;
- la edad media en que las mujeres dejan la familia de origen para ir a vivir solas;[39]

[sólo si la unidad es un municipio urbano]

- el gasto total (en el último año) en platos preparados dividido por el gasto total (en el mismo período) en productos alimenticios no preparados en los supermercados del municipio.

Cuando la unidad es individual, el concepto apropiado es el grado de apoyo a la emancipación femenina. Un abanico de indicadores de éste podría incluir:

- la actitud hacia la asignación de cargos políticos a mujeres;
- la actitud hacia la asignación de cargos religiosos a mujeres;
- la actitud hacia el aborto;
- la actitud hacia los anticonceptivos;
- la actitud hacia las mujeres empresarias;
- el juicio acerca de la conveniencia de que las mujeres trabajen fuera de su casa;
- el juicio acerca de la conveniencia de que las mujeres trabajen de noche.

9.6. La construcción de índices tipológicos con variables categoriales y ordinales

Se dijo varias veces que la intensión de un concepto de propiedad tiene varios aspectos, y que —si no es posible definir operativamente el concepto de forma directa— se necesita encontrar indicadores para los aspectos más importantes, so pena de recoger informaciones sobre un concepto muy diferente del

[39] Los dos últimos indicadores están invertidos, en el sentido de que cuanto más alto es el número de niños menores de cinco años, y más alta la edad en la que las mujeres dejan la familia de origen, más bajo es el nivel de emancipación femenina.

que pensamos. Naturalmente, las informaciones recogidas tienen que ser sintetizadas para reconstituir de alguna manera la unidad del concepto que interesa.

La operación con la que se realiza esta síntesis se llama CONSTRUCCIÓN DE UN ÍNDICE y, tal como se adelantó, es la tarea más delicada de toda la investigación social que sigue una aproximación estándar.

El criterio de construcción de los índices se puede establecer en abstracto, a la hora de diseñar la investigación; pero se realiza cuando los datos sobre los indicadores estén recolectados, y tengamos las variables correspondientes en la matriz. El procedimiento que se sigue es totalmente diferente según se traten las variables como cardinales o no, a punto tal que es preciso presentarlo en párrafos diferentes.[40]

Si no las consideramos cardinales, se atribuye plena autonomía semántica a cada categoría de las variables que van a formar el índice. De esto se sigue que la herramienta para realizar la síntesis es la tabla bivariable,[41] en la que cada categoría de las variables involucradas se representa en una fila o una columna de la tabla, y en cada celda hay un tipo —es decir la combinación de una categoría de la variable en las filas de la tabla con una categoría de la variable en las columnas. El índice así construido se denomina TIPOLÓGICO.

Empecemos con un ejemplo simple, en el que las dos variables involucradas son ordinales (lo que nos ayudará a realizar la síntesis) y tienen pocas categorías (lo que reducirá la complejidad del problema). Imaginemos que la propiedad que nos interesa sea el "capital cultural" (en el sentido de Bourdieu 1979; 1994) que recibió cada individuo durante su formación, y que dos de los indicadores elegidos sean —muy simplemente— el título de estudio del padre y el de la madre, cada uno registrado con cuatro categorías:[42] elemental (*el* en algunas tablas que siguen) / medio (*me*) / medio-superior (*ms*) / superior (*sup*).

Tabla 9.1: *Combinación de dos indicadores y reducción de la tipología*

	TÍTULO DEL PADRE			
TÍTULO DE LA MADRE	ELEMENTAL	MEDIO	MEDIO SUPERIOR	SUPERIOR
elemental	1	2	3	4
medio	2	2	3	4
medio superior	3	3	3	4
superior	4	4	4	5

[40] Hay naturalmente casos mixtos, en los que algunas variables son cardinales o cuasi cardinales y otras no. Un ejemplo se presentará en el apartado siguiente.

[41] Naturalmente hay índices formados con más de dos variables categoriales. Pero también en estos casos, ya que una tabla trivariable no es fácil de concebir ni de representar, en la práctica se opera con una sucesión de tablas bivariables, construyendo índices parciales (véase un ejemplo más adelante).

[42] Se usan categorías generales, sin hacer referencia al ordenamiento escolástico de un país particular. El ajuste de estas categorías generales a la estratificación de títulos de cada país es un típico problema de definición operativa.

La tabla 9.1 presenta los dieciséis tipos resultantes de la combinación (producto lógico) de cada título del padre con cada título de la madre. Estos tipos podrían ser las dieciséis modalidades de la nueva variable. Esta variable deberá ser cruzada con otros indicadores de capital cultural, y por lo tanto se necesitará una secuencia de otras tablas. Dieciséis modalidades en las filas o en las columnas de una tabla son demasiadas; por eso es preciso reducir el número de tipos en la tipología.[43] Si la contribución del padre y la de la madre al capital cultural de un individuo se consideran equivalentes, una manera obvia de reducir esta tipología se muestra en la tabla 9.1:

- el código 1 en la nueva variable se atribuye a todos los que tienen ambos padres con el título elemental;
- el código 2 se atribuye a todos los que tienen al menos uno de los padres con el título medio;
- el código 3 se atribuye a todos los que tienen al menos uno de los padres con el título medio-superior;
- el código 4 se atribuye a todos los que tienen al menos uno de los padres con el título superior;
- el código 5 se da a todos los que tienen ambos padres con el título superior.

Se logra así una nueva variable ordinal, que sintetiza dos indicadores. Naturalmente, se pueden concebir otras formas de realizar esta síntesis, por ejemplo atribuyendo un peso mayor al título del padre (códigos en la tabla 9.2, parte izquierda) o al de la madre (códigos en la tabla 9.2, parte derecha).

Tabla 9.2: *Otras formas de combinar dos indicadores y reducir la tipología resultante*

	TÍTULO DEL PADRE				TÍTULO DEL PADRE			
TÍTULO DE LA MADRE	EL	ME	MS	SUP	EL	ME	MS	SUP
elemental	1	2	3	4	1	1	2	2
medio	1	2	3	4	2	2	3	3
medio superior	2	3	4	5	3	3	4	4
superior	2	3	4	5	4	4	5	5

Antes de mostrar cómo se combina una de las variables así construidas con otro indicador, siguiendo con la construcción del índice de capital cultural, ca-

[43] Dieciséis modalidades en las filas o en las columnas de una tabla, cruzándose con las categorías de otra variable, producen un número demasiado alto de celdas (el número de celdas de una tabla es el producto del número de filas por el número de columnas). La reducción del número de tipos en una tipología, llamado por brevedad "reducción de una tipología", fue teorizada por Hempel y Oppenheim (1936) y por Lazarsfeld (1937). Este último la llamó "reducción del espacio de atributos" por miedo a que el término 'tipología', de noble origen filosófico, fuera rechazado por el mercado norteamericano, en el que estaba desembarcando al huir del nazismo.

be destacar el hecho de que la relación entre el título del padre y el de la madre se puede elegir como uno de los indicadores del nivel de tensión y conflicto potenciales en la familia de procedencia del sujeto, a su vez una posible causa (no un indicador) de su nivel de (in)seguridad. El supuesto es que cuanto mayor sea el desnivel cultural entre los padres, tanto mayores probablemente serán —en el largo plazo— la tensión y el conflicto potenciales entre los dos. En este caso, los códigos a asignar a los tipos en las celdas (véase la tabla 9.3) podrían ser:

- 0 en la diagonal, donde el título es el mismo para ambos padres;
- 1 en las celdas cerca de la diagonal, donde sólo hay un escalón de diferencia entre el título de uno y del otro;
- 2 en las celdas donde hay dos títulos de diferencia;
- 3 en la celda a la extrema derecha alta y a la extrema izquierda baja de la tabla, donde la diferencia es la máxima posible.

Tabla 9.3: *Otras formas de usar las mismas informaciones para crear un indicador de otro concepto*

	TÍTULO DEL PADRE			
TÍTULO DE LA MADRE	ELEMENTAL	MEDIO	MEDIO SUPERIOR	SUPERIOR
elemental	0	1	2	3
medio	1	0	1	2
medio superior	2	1	0	1
superior	3	2	1	0

Veamos ahora cómo se puede proceder en la construcción del índice usando la variable formada por los títulos de los padres en combinación con otro indicador de capital cultural, por ejemplo el tipo de lecturas preferidas por el sujeto durante su adolescencia y su juventud. Imaginemos que las diversas respuestas a estas preguntas hayan sido reagrupadas en cuatro niveles:

1: historietas, crónica deportiva, novelas "rosa" y crónica "social" en diarios y periódicos (hist en la tabla 9.4);
2: cuentos de aventura y de viajes, novelas policíacas, secciones policíacas de diarios y periódicos (av en la tabla 9.4);
3: novelas, *best sellers*, ciencia ficción (nov en la tabla 9.4);
4: ensayos, clásicos de la literatura y del teatro (cl en la tabla 9.4).

En este caso la reducción de la tipología es una tarea delicada, sobre todo porque las categorías en ambas variables incluyen situaciones reales bastante diferentes, que cada investigador (y cada lector) puede interpretar a su manera. En general, los criterios que guían la elección de los tipos que se van a reagrupar (y por consiguiente las categorías de la variable resultante, que puede ser el índice final o una etapa intermedia en su construcción) son los mismos que guían la reducción del número de las clases en una clasificación (véase apartado 8.1):

el primero es la máxima proximidad semántica entre los tipos que se reagrupan (complementada por la máxima distancia semántica entre las categorías resultantes). Naturalmente, el punto de referencia para juzgar todo esto será el concepto general que se quiere operativizar mediante un índice —en este caso el capital cultural. El criterio que ya se ha definido como auxiliar es la oportunidad de evitar desequilibrios excesivos entre las frecuencias de las categorías resultantes, para no perjudicar la sucesiva fase de análisis.

Como siempre, hay que balancear el criterio semántico con el numérico.[44] Este balanceo se realiza más fácil y eficazmente si ya tenemos la tabla ante nuestra vista, es decir, cuando los datos ya han sido recolectados. Por este motivo se aconseja proyectar los índices en general y en abstracto antes de recolectar los datos, pero esperar que ellos sean recolectados antes de encarar los detalles de la construcción.

Desafortunadamente, no se puede cumplir con esta máxima al proponer el ejemplo de reducción que sigue (tabla 9.4). Sin embargo, la propuesta tiene en cuenta una consideración empírica, es decir, el hecho de que muchos entrevistados, para dar la mejor imagen de sí mismos al contestar una pregunta semejante, tienden a indicar —entre todas sus lecturas— aquellas que perciben como más apreciadas culturalmente.

Los tipos que emergen de la reducción podrían ser:

1: capital cultural consistentemente bajo: sujetos cuyos padres no llegaron al nivel medio-superior y que prefieren leer historietas, deportes, cuentos de aventura y de viaje, crónica "social" y novelas "rosas" o policíacas;
2: capital cultural descendente: sujetos cuyos padres alcanzaron el nivel superior y que no llegaron a indicar ensayos o clásicos al contestar la pregunta; sujetos con uno de los padres con instrucción superior y que no llegaron a indicar clásicos o novelas de buen nivel al contestar la pregunta; sujetos con uno de los padres con instrucción media-superior y que prefieren leer historietas, deportes, o crónica "social" y novelas "rosa";
3: capital cultural consistentemente medio: sujetos con uno de los padres con instrucción superior y que no llegaron a indicar ensayos o clásicos al contestar la pregunta; sujetos con uno de los padres con instrucción medio-superior y que prefieren leer cuentos de aventura y de viajes, crónica y novelas de tipos varios (no "rosa"); sujetos con uno de los padres con instrucción media y que prefieren leer novelas de buen nivel, *best sellers*, ciencia ficción;
4: capital cultural ascendente: sujetos sin ningún padre con instrucción superior que indican ensayos o clásicos como su lectura preferida, y sujetos con ambos padres con título elemental que indican ensayos, clásicos, novelas de buen nivel, *best sellers*, ciencia ficción;

[44] Lo destacaba también Galtung (1967: 253).

5: capital cultural consistentemente alto: sujetos con uno o ambos padres con instrucción superior que indican ensayos o clásicos como su lectura preferida.

Tabla 9.4: *Cómo proceder en la construcción de un índice de capital cultural*

	LECTURAS PREFERIDAS			
TÍTULO DE LOS PADRES	HIST	AV	NOV	CL
2 elemental	1	1	4	4
1 medio	1	1	3	4
1 medio superior	2	3	3	4
1 superior	2	2	3	5
2 superior	2	2	2	5

Hasta ahora, sólo se combinaron tres indicadores, y ya el ejemplo puede dar una idea de la cantidad y complejidad de problemas semánticos y de decisiones consiguientes que están involucrados en la construcción de un índice con indicadores cuya definición operativa lleva a una variable categorial u ordinal. Cada vez que se quiera combinar un indicador más, la complejidad semántica se acrecienta de forma casi exponencial, las decisiones involucradas en reducir la tipología son cada vez más cuestionables, y el control intelectual sobre toda la operación y sus resultados disminuye.

Por este motivo, cada vez que se puedan imaginar definiciones operativas que lleven a variables cardinales o cuasicardinales, los investigadores expertos las prefieren. Por otro lado, ser conscientes de las ventajas vinculadas a las variables cardinales no justifica tratar como cardinales variables que ni siquiera son ordinales —como a menudo se encuentra.

9.7. La construcción de índices aditivos con variables cardinales y cuasi cardinales

Construir índices con variables cardinales o cuasi cardinales es más simple que construirlos con variables categoriales u ordinales, y eso por dos motivos. El primero es obvio: la posibilidad de realizar con plena legitimidad[45] operaciones matemáticas con los códigos numéricos que representan los estados en estas variables. El segundo motivo —la autonomía semántica reducida o nula de las categorías (véanse los apartados 8.4 y 8.5)— no es tan obvio, y por eso no se encuentra frecuentemente mencionado en los textos de metodología. Empero, no tiene menor importancia, porque es lo que permite pasar de un tratamiento mediante tablas, donde el foco de la atención se fija en cada celda, a un tratamiento mediante diagramas, donde la atención puede fijarse en el rumbo global de la relación.

[45] En rigor, la legitimidad es plena sólo en el caso de variables cardinales. Como se destacaba en el apartado 8.5, las variables cuasi cardinales presentan diferencias nada despreciables con las cardinales; por eso tratarlas como cardinales no es plenamente legítimo, y se lo hace para aprovechar las grandes ventajas involucradas en este tipo de tratamiento. Con variables ordinales, la distorsión puede ser tan grande que muchos no consideran legítimo tratarlas con técnicas cardinales. Con variables categoriales este tratamiento es un verdadero sinsentido, ya que los códigos numéricos de las categorías son enteramente arbitrarios.

Como veremos en el capítulo 15, esta consideración vale no sólo para la construcción de índices, sino también para el análisis de las relaciones entre variables. Desde el punto de vista técnico, los procedimientos son los mismos: lo que cambia radicalmente es el objetivo del tratamiento. En un caso, estamos construyendo variables de mayor generalidad y porte teórico; en el otro, las variables ya están construidas y estamos explorando un segmento de la tupida red de relaciones entre ellas.

Sin embargo, las operaciones matemáticas necesarias para construir índices son mucho más sencillas de las que se usan en el propio análisis estadístico. A veces son simples sumas de puntajes: el puntaje de cada caso (individuo, agregado territorial u otro) en el índice es igual a la suma de sus puntajes en los indicadores. Por eso los índices de este tipo se denominan ADITIVOS.[46] Naturalmente, este nivel de simplicidad sólo se puede alcanzar si se satisfacen cuatro condiciones: una factual, una numérica y dos semánticas.

La condición factual es que no falten datos en uno o más indicadores: en estas circunstancias el puntaje final de cada caso en el índice debe ser el promedio y no la suma de sus puntajes válidos. Si se sumasen, los puntajes finales de todos los casos que tienen datos faltantes serían indebidamente menores de lo que deberían ser: un sujeto podría resultar menos inteligente, o menos autoritario, o menos conservador, o menos progresista de lo que es realmente. Naturalmente, le corresponde al investigador decidir cuántos datos faltantes se pueden aceptar: si un índice está formado por cinco indicadores y para el sujeto X faltan informaciones acerca de tres de ellos, sería sabio no asignar al sujeto X ningún puntaje en ese índice.

La condición numérica es que todas las variables que se suman tengan la misma extensión de escala, o al menos una extensión muy parecida. Esta condición vale igualmente para índices formados con variables métricas (resultados de una medición), con variables cardinales naturales (resultados de un conteo) y con variables cuasi cardinales (resultados de algunas técnicas de construcción de escalas).

Veamos un ejemplo para cada tipo. Imaginemos que la federación de atletismo de una nación pobre disponga de recursos para enviar a las Olimpíadas sólo un lanzador. Para elegir quién será se decide sumar para cada candidato los resultados que alcanza en los cuatro tipos de lanzamiento (disco, jabalina, martillo y bala) y enviar el que tiene la suma más alta. Pero está claro que de esta forma se perjudica a los lanzadores de bala, cuyo margen con respecto de los no especialistas puede girar en torno de (los) cinco me-

[46] Este rótulo se usa también cuando los puntajes no se suman directamente, sino que antes se someten a las elaboraciones descritas en el texto.

tros, mientras los otros lanzadores pueden tener márgenes de treinta metros y más con respecto de los no especialistas. La solución correcta sería estandarizar los puntajes (véase apartado 8.3), o al menos normalizar[47] cada resultado dividiéndolo por el récord nacional o por el récord mundial en la especialidad respectiva.

Imaginemos ahora que se quiera construir un índice de posesión de bienes muebles duraderos, y para cada sujeto se sumen el número de coches, el número de equipos de TV, de heladeras, de lavarropas, etcétera, hasta al número de equipos de radio y de celulares. Pero está claro que un celular no tiene el mismo valor que un coche: antes de sumar estos artículos se debería ponderar cada uno por su costo específico, o al menos por el costo medio del tipo de artículo.

Por último, imaginemos construir un índice de progresismo. Ya que el progresismo es un concepto multifacético y no se puede medir directamente, hay que encontrar indicadores y operativizarlos mediante las técnicas de construcción de escalas que se ilustran en el apartado 8.5. Por ejemplo, usando los puntajes asignados por cada sujeto a líderes o símbolos progresistas con el "termómetro", su autocolocación con la escala derecha-izquierda, y las respuestas elegidas reaccionando a frases de orientación progresista en una batería de escalas de Likert. Para crear el índice no se podrían sumar directamente estos puntajes,[48] porque los puntajes en los termómetros van de 0 a 100, en la escala derecha-izquierda van de 0 a 10, en las escalas de Likert van de 0 (o 1) a 4 (o 5). Antes de sumar se deberá estandarizar los puntajes, o al menos normalizarlos dividiendo cada uno por el máximo puntaje previsto para la escala relativa.

Algunos ejemplos ya presentados sirven también para ilustrar la primera condición semántica. Imaginemos que la federación de atletismo también quisiera enviar a las Olimpíadas un decatleta. Es sabido que las pruebas del decatlón incluyen tres lanzamientos, tres saltos y cuatro carreras. En este caso, no basta con normalizar los puntajes antes de sumarlos. Esto porque en los saltos y lanzamientos cuanto más alta es la cifra, mayor es el mérito; para las carreras, vale obviamente lo contrario. Por eso, se debe invertir la dirección de los puntajes relativos a las carreras —y esto es exactamente lo que hace la federación mundial de atletismo en su tabla de puntajes para el decatlón.

El mismo problema semántico se presenta con el índice de progresismo. Si también se quiere insertar en el índice los puntajes de agrado asignados con el termómetro a líderes y símbolos conservadores, y las respuestas a frases conservadoras en una batería de escalas de Likert, puede hacérselo sólo después de invertir la dirección de los puntajes.

La segunda condición semántica de la cual se hablaba es más compleja. Hasta ahora se presentaron operaciones (estandarización, normalización, inversión de dirección semántica) que tenían el fin de hacer efectivamente equivalentes los indicadores antes de combinarlos en un índice, en el supuesto de que el investigador los considerara todos igualmente válidos.[49] Pero puede pasar que un investigador considere algunos más válidos que otros, por motivos que pueden radicar en su conocimiento directo del problema y/o del ámbito investigado, en sus experiencias previas de investigación, en la literatura sobre el tema. Si tiene buenos motivos para creer que algunos indicadores son más válidos que otros, tiene que incorporar esta convicción en la formación del índice, sin temor a no ser "objetivo". Más allá del hecho de que la elección de indicadores no puede de ningún modo ser "objetiva", se debe agregar que considerar igualmente válidos todos los indicadores es un supuesto, exactamente como lo es lo contrario.

[47] La estandarización es sólo una forma sofisticada de normalización.

[48] Sin embargo, hay quien lo hizo —por ejemplo, el politólogo norteamericano Smith (1969)— sin darse cuenta de que de esta forma estaba considerando cada termómetro veinte veces más importante que cada escala Likert.

Para ponderar la contribución de cada indicador a la suma que va a proporcionar a cada caso su puntaje en el índice, el investigador tiene dos caminos. Puede guiarse por sus valoraciones semánticas, multiplicando los puntajes en los indicadores que considera más válidos por un coeficiente mayor de 1 (puede ser 1,2 o 1,5 o algo parecido) y los indicadores que considera menos válidos por un coeficiente menor de 1 (puede ser 0,7 o 0,8 o algo parecido). O puede aprovechar las posibilidades que le ofrece la matriz de los datos, sometiendo la canasta de las variables que operativizan posibles indicadores del concepto que interesa a un análisis de componentes principales.[50]

Ahora basta con decir que el análisis de componentes principales tiene tres funciones, a las que corresponden tres fases en el procedimiento:

1) seleccionar en la canasta inicial de posibles indicadores los que resultan tener una estricta relación empírica con los otros, y por lo tanto se puede suponer que tengan una apreciable relación semántica con el concepto que interesa;
2) estimar la fuerza de esta relación semántica (es decir la validez) entre el concepto y cada uno de los indicadores seleccionados;
3) basándose en esta estimación, proporcionar un coeficiente para ponderar los puntajes de cada indicador teniendo en cuenta las posibles superposiciones semánticas, es decir el hecho de que dos o más indicadores pueden representar aspectos muy cercanos de la intensión del concepto.

En este caso, el índice final no es otra cosa que la suma de los puntajes de cada individuo en los indicadores seleccionados en la fase (1); antes de sumarlos, los puntajes son multiplicados por el correspondiente coeficiente calculado en la fase (3).

[49] Sobre el concepto de validez, véase el apartado 9.4.

[50] El análisis de componentes principales es preferible al más antiguo y más difundido análisis factorial, porque no realiza manipulaciones arbitrarias de las matrices de los datos antes de someterlas a tratamiento con el álgebra matricial. Una reseña de otras técnicas antiguas para ponderar indicadores basándose en sus relaciones con otras variables en la matriz está en Edgerton y Kolbe (1936).

Capítulo 10

LA OBSERVACIÓN

10.1. Introducción

La observación es literalmente el "acto y efecto de observar",[1] que a su vez significa mirar o examinar con atención. A partir de esta primera aproximación queda en evidencia que la observación no es un patrimonio exclusivo de las prácticas científicas de investigación; se trata de una actividad con presencia constante y recurrente en la vida cotidiana. En este sentido, podría concebírsela como un medio de conocimiento y un prerrequisito del relacionamiento con los demás; independientemente de los distintos enfoques con los que se ha tratado de dar cuenta de la vida social y la sociabilidad, la observación —en sentido amplio— ha jugado siempre un rol destacado.

En el ámbito de la ciencia, la observación es, por definición, el modo de establecer algún tipo de contacto empírico con los objetos/sujetos/situaciones de interés a los fines de su descripción, explicación, comprensión.

Estas consideraciones genéricas que hemos hecho hasta aquí dan lugar a dos cuestiones problemáticas cuyo abordaje articulará gran parte del resto de este capítulo:

a) Si la observación no es patrimonio exclusivo de la ciencia, ¿en qué se diferencia la observación en la vida cotidiana de la observación científica?

b) Si en la ciencia la observación es el camino privilegiado para construir conocimiento empírico, ¿hay distintas formas de observación? ¿En qué se diferencian? ¿Cómo se las puede clasificar?

[1] Definición del *Diccionario de la lengua española* de la Real Academia Española, vigésima segunda edición, 2001.

10.2. La observación en la vida cotidiana y en la ciencia

Tal como plantea Valles (1997), la distinción entre observación científica y observación común es una de las cuestiones centrales que ha caracterizado a la reflexión metodológica sobre este tema. En términos generales las relaciones entre observación en la ciencia y en la vida cotidiana han sido entendidas recurriendo a las nociones de ruptura o de continuidad.

La idea de ruptura destaca el quiebre entre ambos mundos, y se inscribe en el marco de una concepción que le asigna a la ciencia un lugar privilegiado, inmune a las "contaminaciones" del conocimiento ordinario que caracteriza a la vida cotidiana. Ésta es una noción de larga tradición en el pensamiento occidental, desde que algunos de los más influyentes filósofos griegos —en especial Platón— definieron a la ciencia como conocimiento cierto, claro (επιστημη) por oposición al conocimiento vulgar o mera opinión (δοξα).

En contraste, en tiempos más recientes se ha tendido a favorecer la idea de continuidad entre ciencia y vida cotidiana, y en este sentido se ha destacado que las competencias observacionales desarrolladas en la vida social, junto con un esfuerzo reflexivo y crítico en torno de sus características, su significado e implicaciones, constituye un fundamento de notable importancia no sólo para aprender y comprender el significado de la observación científica, sino también para llevarla a cabo de manera apropiada. Como expresan Ruiz Olabuénaga e Ispizua (1989: 79), "la observación común y generalizada puede transformarse en una poderosa herramienta de investigación social y técnica científica".

En esta línea, la observación científica puede concebirse —retomando las palabras de Cardoso de Oliveira (1996) para referirse al trabajo antropológico— como una forma de observación disciplinada, y esto en un doble sentido: disciplinada en cuanto caracterizada por la sistematicidad y la constancia, una práctica que se atiene a ciertas reglas y procedimientos, y disciplinada en la medida en que está orientada teórica y metodológicamente por las disciplinas científicas, en el marco de las cuales adquiere un sentido específico que a su vez es producto de consensos más o menos generalizados acerca de las reglas, procedimientos de acción y perspectivas teóricas que la guían.

10.3. Distintas formas de observación científica

Se dijo más arriba que todos los métodos de investigación empírica son en algún sentido observacionales. Rebello (1977), para dar cuenta de su amplitud en la práctica científica, define a la observación como la más primitiva y a la vez la más actual de las técnicas de investigación.

Una primera y elemental distinción entre las formas de observación se puede establecer a partir de la coincidencia (o no) entre el sujeto observador y el objeto/sujeto observado. Cuando ambos coinciden, se está en presencia de lo que se conoce como autoobservación, que será más o menos introspectiva según el grado de profundidad del análisis de la propia personalidad y comportamiento. En los otros casos (es decir, de no coincidencia entre observador y

observado) nos referimos a la observación en sentido genérico, que a su vez puede dividirse en una multiplicidad de variedades de acuerdo con diferentes criterios. Por otra parte, al empleo de distintos *fundamenta divisionis* para operar la clasificación de la observación, se suma el hecho de que algunas de ellas se realizan siguiendo un esquema jerárquico, en el que una clase es a su vez dividida en varias subclases, dando lugar a taxonomías.

En sentido amplio suele distinguirse entre observación directa e indirecta. Esta última se da cuando el investigador adquiere conocimiento del objeto de estudio a través de observaciones previas realizadas por otras personas; por ejemplo, cuando recurre a datos secundarios. La primera se refiere, en cambio, a los casos en que el investigador se pone en contacto directo con el fenómeno de interés, y esto se puede hacer de manera controlada (por ejemplo, a través de experimentos) o no controlada, cuando se observa la situación en su ambiente natural, sin alterar el contexto ni recrearlo artificialmente. También es frecuente referirse a la observación controlada como observación de laboratorio, y a la no controlada como observación de campo o naturalista. A su vez, esta última puede ser participante o no participante, según se observe el fenómeno "desde adentro" (involucrándose e interviniendo en la situación objeto de estudio) o "desde afuera" (adquiriendo una actitud externa de mera contemplación).

Otras formas de definir a la observación incluyen a la estructurada o no estructurada, según el grado de sistematicidad y de delimitación previa de qué/quién, cómo y cuándo será observado; retrospectiva, cuando se investigan fenómenos del pasado (recurriendo a documentos, entrevistas, relatos de vida, etcétera) y mediada o no mediada, según se apele a instrumentos técnicos de observación (telescopio, microscopio, etcétera) o simplemente a los sentidos, sin la ayuda de tales dispositivos técnicos. Debe quedar claro que no nos referimos en este caso a los medios para el registro de las observaciones (grabadores, cámaras fotográficas y de video o instrumentos de medición diversos) sino a aquellos que hacen posible un tipo determinado de observación. En este aspecto, se puede distinguir entre los instrumentos que amplían la capacidad de los sentidos y, en un modo más general, aquellos que hacen posible acceder a través de ellos a la observación de un espacio/objeto dado. Si bien en la investigación social el uso de medios que amplían la capacidad de los sentidos es muy limitado, éste no es el caso del segundo tipo de mediación. En efecto, y desde hace aproximadamente unos quince años, se viene difundiendo la idea de observación virtualizada, cuando ésta —como afirman Ardèvol *et al.* (2003)— se vehiculiza a través de una herramienta informática. El foco de interés de tales observaciones son las interacciones que se desarrollan en el ciberespacio: tómese por caso el chat, que para muchos constituye una nueva forma de sociabilidad (véase por ejemplo Jones 1995, 1997; Turkle 1995). Todas las conversaciones, interacciones, etcétera, que se observan y de las que se es partícipe quedan completamente condicionadas por aspectos técnicos: los *softwares* que se utilizan para acceder a los chats y el *hardware* necesario para ejecutarlos (Ardèvol *et al.* 2003).

Habiendo hecho las consideraciones precedentes, resulta pertinente consignar que en este capítulo nos interesa abordar en detalle las formas de obser-

vación directa, y en especial aquellas que se conocen como no controladas (de campo, naturalistas), intensivas, participantes y no mediadas.

Gold (1958), siguiendo a Buford Junker, de cuyo equipo de investigación había sido miembro, profundizó en la caracterización de los tipos ideales de roles de trabajo de campo que el cientista social puede adoptar a partir de la observación y la participación. Su tipología se basa en dos criterios: el grado de participación (involucramiento versus distanciamiento) y el grado de revelación de la actividad de observación. De aquí surgen los ya clásicos cuatro tipos teóricos: participante completo, participante-como-observador, observador-como-participante, observador completo.

En el caso del participante completo, la verdadera identidad y propósitos del investigador se mantienen ocultos a aquellos que se observa; el investigador interactúa de la manera más natural posible en las situaciones que le interesan, aprendiendo a desarrollar roles y tareas de la experiencia cotidiana de los observados como si fuera uno de ellos. En este sentido, es análogo a lo que se conoce habitualmente como observador encubierto (véase apartado 10.4).

El participante-como-observador es un rol de trabajo de campo con características similares a las del anterior, pero se diferencia en que tanto investigador como sujeto observado son conscientes de que la suya es una relación de campo. Gold (1958) indica que éste es el típico estilo de trabajo de campo en el que el observador establece relaciones con informantes a lo largo del tiempo, y tiene la oportunidad de participar en las situaciones investigadas. Como veremos más adelante, ésta es la modalidad de observación participante que más se ajusta a la idea actual que se tiene de ella en la antropología y la sociología.

El observador-como-participante es el típico rol de quien limita el trabajo de campo a una entrevista. En ella la actividad de observación es más bien formal, ya que no se dan las condiciones para una observación y participación informales a lo largo del tiempo. De este tema nos ocuparemos en el capítulo 12.

El observador completo, finalmente, es aquel que no interactúa socialmente con los sujetos investigados; se limita a observar intentando que su actividad pase inadvertida por parte de ellos.

Burgess (1984) indica que los cuatro tipos ideales planteados por Gold no deben entenderse como compartimientos estancos. Por el contrario, plantea que existe dinamismo e interconexión entre ellos, y que incluso a lo largo de un mismo estudio el investigador puede adoptar distintos roles de observación y participación. Además, se debe tener en cuenta que el desarrollo de un rol determinado no se da unilateralmante, como si esto fuese patrimonio exclusivo del investigador. Por el contrario, éste también es objeto de observación, y desde perspectivas diversas, hecho que deriva en la asignación de roles por parte de los sujetos investigados. En este sentido, Guber (1991/2005: 87) señala que "el investigador no es el único estratega"; en el proceso de conocimiento se dará una "interacción, diferenciación y reciprocidad entre la reflexividad del sujeto cognoscente —sentido común, teoría, modelo explicativo de conexiones tendenciales— y la de los actores o sujetos/objetos de investigación".

En vez de pensar en tipos cerrados de observación y participación, muchos autores favorecen la idea de un *continuum* de diversas combinaciones posibles entre los polos de participación y observación puras. Spradley (1980), por ejemplo, sostiene que el grado de participación puede variar desde la no participación (cuando se observa desde fuera del escenario de investigación) hasta la participación completa, pasando por la participación pasiva y la participación moderada. Por otra parte, se puede afirmar en cierto sentido que toda participación presupone observación, y que toda observación —salvo que el observador esté físicamente oculto a los ojos de los sujetos observados— implica participación, aun en un sentido muy débil, en la medida en que de algún modo su presencia incidirá en la situación que observe.

10.4. La observación participante

Como ya se ha aludido en párrafos precedentes, esta estrategia consiste en presenciar de manera directa el fenómeno estudiado en su "ambiente natural", sin manipularlo. Kawulich (2005, par. 81) afirma que "la OBSERVACIÓN PARTICIPANTE conlleva el involucramiento del investigador en una variedad de actividades, y por un período prolongado de tiempo, con el fin de observar a los miembros de una cultura en su vida cotidiana y participar en sus actividades facilitando una mejor comprensión de los mismos".[2] Se trata de un proceso que "implica el acceso a la comunidad de interés, la selección de porteros e informantes clave, la participación en todas aquellas actividades que los miembros de la comunidad permitan, la clarificación de los hallazgos a través de controles con algunos de los miembros, de entrevistas formales y de conversaciones informales, y el registro de notas organizadas y estructuradas que faciliten el desarrollo de una narración que explique diversos aspectos de esa cultura".

Osorio (1999) distingue entre modos pasivos y activos de observación y de participación. Los modos pasivos de observación implican el mirar (en cuanto registro visual que permite describir lugares, ceremonias y personas) y el escuchar, mientras que el modo activo es el conversar. Los modos pasivos de participación son el estar (que implica tan sólo la presencia física del investigador en el terreno) y el hacer (que involucra la realización de alguna actividad física); el modo activo es el compartir. A partir de estas consideraciones distingue entre la observación participante en un sentido débil (modos pasivos) y fuerte (modos activos). Osorio defiende las ventajas de la forma débil como medio para investigar y conocer en las ciencias sociales porque a su juicio, en el sentido fuerte, sólo se gana familiaridad pero no conocimiento. Además, este último conlleva un estar de acuerdo, un hacer propios los significados de la cultura extraña, acercándose a lo que en términos clásicos se conoce como "convertirse en un nativo". Spittler (2001) aboga, en cambio, por un aumento de la participación, una forma radicalizada de observación participante que llama "participación densa" y que se orienta a la maximización de lo que considera como puntos fuertes de esta forma de investigar.

[2] El énfasis es nuestro.

La observación participante también puede adquirir diferentes sentidos según cuál sea la relación de membresía del investigador con respecto al grupo que se estudia; el observador puede ser miembro (nativo o no nativo) del mismo, puede hacerse pasar por miembro, o desarrollar un rol de observador participante en el grupo en el sentido del segundo de los tipos propuestos por Gold (1958).

En cuanto miembros de una comunidad, todos somos observadores participantes de ésta, y éste fue justamente el sentido de la expresión en sus orígenes: tal como sostiene Hammersley (1989), cuando fue introducida en 1924 hacía referencia al observador "nativo" al que luego recurría el investigador para obtener información sobre la situación estudiada, convirtiéndolo en lo que hoy llamaríamos un INFORMANTE CLAVE.

La observación de la propia comunidad exige un deliberado trabajo de extrañamiento/distanciamiento que permita desnaturalizar prácticas constitutivas de la cultura a la que se pertenece y que el observador ha adquirido a través de procesos de socialización. La tradición de las ciencias sociales ha tendido a pensar al observador como un "extraño que está dentro" de la comunidad, que adquiere ciertas competencias sociales necesarias para participar en ella, pero que no debe hacer propias las ideas y valores característicos de ésta. En otras palabras, no debería ser un miembro nativo ni convertirse en tal. En torno de este punto, sin embargo, hay ciertos límites sutiles que en las últimas décadas han suscitado importantes debates epistemológicos, metodológicos, e incluso éticos y políticos.

A partir de los postulados de Winch (1964), por ejemplo, se han discutido los límites de la interpretación de una cultura extraña sin que medien procesos de resocialización en el juego de lenguaje que esa cultura implica. Asimismo, se ha puesto en cuestión la posibilidad de "traducción" de una cultura a los términos de otra y a partir de una racionalidad que le es ajena. La supuesta objetividad garantizada por la distancia del ser no nativo puede convertirse también en un obstáculo para acceder al sentido que ciertas prácticas sociales y culturales tienen en el marco de una comunidad que nos es extraña. En contraste, el famoso antropólogo y sociólogo Srinivas (1966, 1998), por ejemplo, ha defendido la ventaja de ser parte de la cultura india para poder comprender el sistema de castas, o el cambio social y cultural en esa nación —lo que constituyó el interés central de sus múltiples investigaciones.

Sin embargo —como ya se ha apuntado—, la visión canónica ha tendido a favorecer las ventajas del observador profesional no nativo, que si bien no debe ser necesariamente un extraño en sentido fuerte, debe poder experimentar la situación de interés desde la doble condición de miembro y extraño. Cabe hacer notar en este sentido que los límites entre el ser miembro o no del grupo observado no son rígidos, y desde que la observación participante se aplica comúnmente al estudio de diversas situaciones y grupos de la sociedad urbana contemporánea, la distinción tajante entre nativo y no nativo —más evidente en el caso del etnógrafo europeo que estudiaba por ejemplo tribus africanas— se ha vuelto más difusa.

Valles (1997: 150), basándose en Spradley (1980), resume las diferencias que se encuentran entre el rol técnico de un observador participante y el papel profano de participación típico de un miembro nativo del grupo. Entre ellas

destaca los contrastes de propósitos (implicarse y observar versus participar "naturalmente"), de los modos de atención (atención incrementada versus desatención selectiva), de los ángulos de observación (abierto versus cerrado), de las formas de experiencia (desde adentro y desde afuera versus sólo desde adentro), de introspección (aplicada versus natural) y de registro (registro sistemático versus no registro).

Cuando el observador se hace pasar por miembro del grupo estamos en presencia de lo que se conoce como observación encubierta (participante completo en la tipología de Gold). Si bien se tiende a desaconsejar esta forma de investigación por razones éticas (Burgess 1984), hay casos en que podría ser la única forma de conocer un fenómeno sin que la conspicua presencia del observador incida decisivamente en la situación observada. Ejemplo de esto es el estudio de Homan (1978) sobre sectas pentecostales, que luego dio lugar a un conocido artículo (1980) sobre los aspectos éticos de los métodos encubiertos en el que justifica su empleo sobre bases pragmáticas, con el fin de evitar la reactividad (véase apartado 10.5) y acceder a aspectos ocultos o secretos. Homan reconoce, sin embargo, que esto puede ser psicológicamente perjudicial para el investigador y dejar huellas duraderas en su personalidad.

En la actualidad —como se dijo más arriba—, la observación participante tiende a presentarse, en el marco de una investigación, como el desarrollo de un rol que se caracteriza por la presencia física del investigador en el terreno; se trata de alguien que se une temporalmente a un grupo con fines científicos y que no oculta su función de investigador. La revelación de su identidad se da al menos ante los "porteros" (*gatekeepers*) que regulan el acceso a ciertos espacios observacionales y están en condiciones de conceder el "permiso" para que el investigador realice allí su tarea.

Para muchos autores, este sentido actual de la observación participante implica la COMPRENSIÓN (Bernard 1994) a través de un acceso privilegiado a los significados que los actores construyen (y le asignan) a su mundo (Platt, 1983). Sin embargo, como la misma Platt (1985) ha sugerido, esta supuesta mutua implicación entre la observación participante como técnica de investigación (y los métodos llamados cualitativos en general) y las ideas interpretativas como su fundamento teórico-filosófico es relativamente reciente. Las prácticas tempranas de observación participante de la antropología y la sociología europea y norteamericana (fines de siglo XIX y principios del siglo XX), incluso en los casos más conocidos como los de la Escuela de Chicago y la antropología funcionalista inglesa (Malinowski, entre otros), no estaban imbuidas de tales bases teóricas y epistemológicas. Piovani *et al.* (2006), al reconstruir históricamente los procesos de producción y reproducción de sentidos en torno de lo cuantitativo y lo cualitativo en las ciencias sociales, sugieren que fue básicamente en el contexto de lo que Giddens (1979) llama disolución del consenso ortodoxo de las ciencias sociales (a partir de la década de 1960) cuando una serie de prácticas de investigación ya existentes —entre ellas la observación participante—, que incluso en muchos casos habían surgido desde la lógica canónica de "recolección" de datos "objetivos", fueron resignificadas desde tradiciones teóricas no ortodoxas (hermenéutica, fenomenología, interpretativismo) para dar susten-

to a una alternativa de investigación sociológica centrada en el acceso interpretativo al mundo de la vida social (véase apartado 2.3).

Volviendo a la observación participante en cuanto rol de campo desarrollado por el investigador, es importante recordar, como se dijo antes, que esto no se puede pensar como una cuestión lineal y autónomamente definida por parte del investigador, independientemente del tipo de relaciones que logre establecer con los actores en el campo y de las expectativas mutuas que se generen y construyan a lo largo del tiempo. En este sentido, cobran especial importancia las características personales del observador. En efecto, muchos autores —entre ellos Burgess (1984) y Kawulich (2005)— han mostrado que el género, la edad, la raza y la clase social del investigador, entre otros factores, inciden en la práctica observacional, en las situaciones que se logra observar y en el tipo de relación de campo que se entabla con los distintos actores.

10.5. Usos de la observación participante en la investigación social. Ventajas y límites

En los manuales de metodología actuales es frecuente recurrir a una presentación esquemática de los usos, ventajas y limitaciones de las diferentes técnicas de investigación. A continuación reseñamos algunas de ellas en el caso de la observación participante, añadiendo oportunamente comentarios ampliatorios y observaciones críticas.

En un sentido genérico, ha tendido a destacarse la utilidad de la observación participante para realizar estudios exploratorios y descriptivos, y su muy limitada capacidad como instrumento de la investigación explicativa, ya sea que se trate de construir teorías o de poner a prueba hipótesis científicas (véase por ejemplo Singleton *et al.* 1993). Sin embargo, esta visión estereotipada que restringe el uso de la observación participante a fases preparatorias —como mera "disparadora" de hipótesis que deberían dar lugar a estudios sucesivos más acordes a la lógica de la contrastación— ha sido desafiada por muchos científicos sociales en las últimas décadas. Glaser y Strauss (1967), por ejemplo, han tratado de mostrar su utilidad (y la de otras técnicas no estándar) para la construcción de teorías formales y sustantivas.

Jorgensen (1989)[3] indica que la observación participante es especialmente adecuada cuando se sabe poco del fenómeno de interés (aspecto exploratorio al que acabamos de aludir), cuando se desea conocer subgrupos cuyos puntos de vista difieren significativamente de los vigentes en una sociedad (minorías, subculturas, contraculturas, etcétera) y cuando la cuestión bajo estudio se oculta a la luz pública (adicciones, criminalidad, sectas religiosas, etcétera). Singleton *et al.* (1993) destacan también su utilidad para dar cuenta de situaciones cambiantes y dinámicas, o cuando resulta imprescindible a los fines del estudio preservar la interrelación entre sujeto y situación. Asimismo, indican que puede ser un buen complemento de otras formas de investigación, no sólo

[3] Citado por Valles (1997: 160).

como proveedora de sugerentes líneas para estudios de tipo explicativo —como se indicó más arriba—, sino también para dotar de profundidad y sentido a los resultados de encuestas y experimentos.

Entre las principales ventajas de la observación participante, Valles (1997) destaca:

a) que en cuanto modo directo de investigación, permite acceder a las situaciones investigadas en toda su complejidad y en el momento en que los acontecimientos suceden, sin artificios ni simplificaciones;
b) que permite acercarse al punto de vista de los actores investigados;
c) que es flexible, posibilitando de este modo redefinir aspectos significativos a lo largo del proceso de investigación.

Entre sus límites señala que hay importantes cuestiones de interés para las ciencias sociales que no son pasibles de observación directa, o al menos de observación participante. Piénsese por ejemplo en los acontecimientos pasados, a los que sólo podemos acceder a través de documentos escritos, del relato de los actores o de otro tipo de fuentes. Valles (1997) también menciona las limitaciones de la observación participante para estudiar fenómenos que tienen una raíz psicológica profunda y que en consecuencia requieren de métodos indirectos (por ejemplo, tests proyectivos), y para abordar cuestiones con gran dispersión poblacional y/o geográfica. Singleton *et al.* (1993) destacan como negativo el tiempo que insume la observación participante, su necesaria limitación a uno o unos pocos escenarios, su ineficacia para obtener cierto tipo de datos y su alta dependencia de las cualidades personales del investigador. Burgess (1984), por su parte, incluye dentro de los límites aspectos de naturaleza ética: ¿en qué medida puede el investigador inmiscuirse legítimamente en la intimidad y privacidad de las personas?[4]

Pero las ventajas y los límites no deben entenderse como absolutos sino en relación con otros métodos observacionales y teniendo en cuenta el objeto de interés de una investigación. Los manuales destacan habitualmente con respecto a los experimentos y las encuestas las debilidades de la observación participante derivadas de la variabilidad del observador —en cuanto "instrumento" humano de recolección de información—, que acarrea potenciales problemas de fiabilidad. También tienden a cuestionarse las posibilidades de generalización a partir del trabajo de campo. Se valora, en cambio, la menor reactividad, el acceso a datos de primera mano y la riqueza de las informaciones obtenidas.

La REACTIVIDAD es la tendencia de los actores a modificar sus comportamientos habituales como consecuencia de saberse observados. La idea más difundida en la metodología de las ciencias sociales es que la observación participante genera menos reactividad que las otras técnicas de recolección. Sin embargo, esta afirmación puede generar confusiones; en realidad, la observa-

[4] A esta cuestión debe sumarse otro importante interrogante de tipo ético: ¿cómo debe actuar el investigador en caso de obtener acceso a información sobre conductas delictivas a partir del trabajo de campo? (véanse Burgess 1984, Singleton *et al.* 1993).

ción participante es hiperreactiva, sólo que por sus características conlleva la oportunidad de que un observador experimentado logre la confianza y empatía suficientes como para diluir los efectos de la reactividad a lo largo del tiempo. Que la observación participante tenga efectos reactivos se puede comprobar en distintas situaciones de la vida cotidiana: cuando una clase de la escuela es visitada por una supervisora del Ministerio de Educación, o cuando un invitado especial y desconocido participa de una cena familiar en el hogar. En estos casos, está claro que muchos de los comportamientos y discursos diferirán del "normal" discurrir de los acontecimientos y se "adaptarán" a la nueva situación en función, por ejemplo, de las expectativas que estos "extraños" generen entre alumnos y maestros, o entre los miembros de la familia. La diferencia de estas situaciones con la observación de campo típica de las ciencias sociales es que esta última permitirá, con el transcurso del tiempo, lograr la confianza (que a veces es tan sólo acostumbramiento) como para que muchas de las conductas "adaptadas" a la presencia del "extraño" pierdan sentido y se vuelva progresivamente al estado habitual de cosas.

Para maximizar las ventajas y restringir los límites de la observación participante, especialmente en términos de fiabilidad, se suele sugerir una serie de estrategias y recomendaciones (véanse por ejemplo Denzin 1970, Burgess 1984, Valles 1997):

- Analizar críticamente la información provista por informantes clave, prestando especial atención a su estatus en el marco de la situación/comunidad/grupo observado. Kawulich (2005), recordando la controversia Mead-Freeman,[5] advierte que los investigadores pueden entender de manera distinta lo que observan según quién/quiénes sean sus informantes clave. En muchas ocasiones, el solo hecho de que alguien esté dispuesto a colaborar con el investigador en calidad de informante lo ubica como un actor atípico o marginal. Guber (1991/2005: 127), por otra parte, plantea que lo apropiado es considerar al informante como un sujeto de investigación, cuya presencia "no garantiza por sí sola un acceso cabal y completo al objeto". Dado que no existe una perspectiva unilateral de conocimiento y que el informante provee una visión mediada por sus preinterpretaciones, propone someter sus relatos a los mismos procesos de construcción y deconstrucción a los que somete su propio conocimiento antropológico.
- Llevar un registro de lo que se percibe como efecto reactivo, en particular en lo que se conoce como notas metodológicas (véase apartado 10.6). Denzin (1970) sugiere incluso que el registro de efectos reactivos se puede realizar también a través de las percepciones de los informantes clave.
- Realizar cruces y controles con los resultados de observaciones realizadas por otros miembros del equipo (en el caso de una investigación de la que participen varios investigadores) y complementar las observaciones con entrevistas y documentos de distinta naturaleza, especialmente con lo que Plummer (1983) llama documentos de vida (*documents of life*). A veces también traducidos como documentos personales (fotografías, diarios, cartas, filmaciones familiares, etcétera), ellos pueden, según Plummer, facilitar un tipo de comprensión profunda de las vidas humanas concretas difícil de alcanzar a través de las fuentes tradicionales de la investigación social.

[5] La conocida antropóloga norteamericana Margaret Mead, discípula de Franz Boas, publicó una monografía en 1928 con los resultados de sus investigaciones con niñas y adolescentes en Samoa. En 1983, el antropólogo neocelandés Derek Freeman publicó un libro en el que se ponen en tela de juicio las conclusiones de Mead. Basándose en su propia experiencia de trabajo de campo en Samoa y en entrevistas realizadas con algunas de la informantes de Mead, Freeman argumentó, entre otras cosas, que ellas habían mentido en sus relatos sobre la conducta sexual. En consecuencia, a los ojos de Freeman las conclusiones del trabajo de Mead eran erróneas (especialmente en lo que respecta a la institución de la virginidad como limitada únicamente a las adolescentes de alto estatus social).

Las críticas más frecuentes a la observación participante remiten al clásico lenguaje de la fiabilidad y la validez. Sykes (1991), en respuesta, afirma que se ha hecho un esfuerzo demasiado grande para justificar la investigación no estándar en términos que no le son apropiados. Por su parte, Adler y Adler (1994) indican que los problemas de fiabilidad y validez en la observación participante tienen solución si se contrastan las perspectivas de varios observadores (con diferentes *backgrounds*), alentando una suerte de *inter-observer agreement*; si se hace una búsqueda deliberada de casos negativos que contradigan las inferencias preliminares del observador, y si se lleva un registro de la variabilidad de los fenómenos en espacio y tiempo a través de observaciones sistemáticas. A su juicio, el trabajo del observador debe reflejar autenticidad y verosimilitud, abonando de este modo la idea de que el trabajo no estándar no debe ser juzgado recurriendo a los términos científicos tradicionales sino a términos alternativos. En este mismo sentido, y como respuesta a las críticas acerca de los límites de la generalización de los resultados obtenidos por medio de la observación participante, Patton (1986) ha propuesto la idea de extrapolación razonable, una especulación modesta acerca de la posible aplicabilidad de los hallazgos a situaciones similares, aunque no idénticas.

10.6. El registro de la información en la observación participante

En primer lugar hay que tener en cuenta que toda observación es focalizada, y que el foco de ésta puede cubrir distinto tipo de espacios, eventos, actividades, conductas, actores, manifestaciones de sentimientos, objetos, momentos, etcétera (Burgess 1984). Para llevar un registro de estas observaciones se recurre a lo que habitualmente se conoce como notas de campo. Bajo este rótulo se incluyen distintas modalidades de registro escrito de la información recolectada a partir del trabajo de campo.

La forma más común de distinción del registro informacional escrito involucra a las notas observacionales, teóricas y metodológicas (Schatzman y Strauss 1973), también llamadas sustantivas, analíticas y metodológicas (Burgess 1984).

Las primeras se refieren al registro descriptivo de lo observado: quién, cuándo, cómo, qué, dónde, etcétera. En ellas es importante rescatar el lenguaje nativo, es decir, la terminología usada por los actores espontáneamente en las situaciones observadas, y hacer un esfuerzo explícito por mantener limitadas al máximo posible las interpretaciones del investigador a partir de lo observado. Pero dado que toda observación conlleva un ejercicio analítico más o menos consciente, se recomienda también realizar notas teóricas o analíticas. Éstas se refieren a los intentos preliminares del observador por dar sentido a las observaciones de campo: se trata de conjeturas, inferencias, etcétera, en las que obviamente se introduce el lenguaje experto de la disciplina. Las notas metodológicas, finalmente, pueden pensarse como apuntes descriptivos acerca del investigador y el proceso de investigación (Schatzman y Strauss 1973): ¿qué decisiones se tomaron en el campo?, ¿con qué fines?, ¿con qué consecuencias? En la medida que tanto el investigador como los sujetos y las situaciones observadas cambian en el proceso de trabajo de campo, es muy importante que las notas queden debidamente identificadas en cuanto a fechas, lugares y actores involucrados.

Spradley (1980), por su parte, distingue entre notas condensadas, que se realizan mientras se llevan a cabo las observaciones, y notas expandidas, que se realizan fuera del campo a partir de las anteriores, y que implican un *racconto* más detallado de lo observado. Asimismo, recomienda llevar un diario de campo, una suerte de diario personal en el sentido estricto, en el que se registran las emociones e impresiones suscitadas a partir del trabajo de campo.

Se ha sugerido más arriba que toda observación implica análisis, y conlleva además un esfuerzo de registro escrito. En este sentido, se recordará también lo expresado en el apartado 4.2 con respecto a la relativa simultaneidad de las tareas de selección, recolección y análisis en este tipo de investigación que hemos definido como no estándar. La concepción tradicional del trabajo de campo, en cambio, ha tendido a diferenciarlas más rígidamente y a ligarlas con fases consecutivas de la investigación, siguiendo "los parámetros del naturalismo: captación inmediata de 'lo real', 'recolección de datos' y posterior análisis en gabinete" (Guber 1991/2005: 251).

A pesar de las críticas que se le han hecho a esta visión canónica, y que compartirmos, creemos igualmente útil distinguir —siguiendo a Geertz (1988)— entre un "estar allí", en la situación de campo (con el énfasis puesto en la observación participante, con todo lo que ésta implica aun en términos de análisis y escritura) y un "estar aquí", en el espacio institucional de trabajo del investigador —que no necesariamente tiene que ser una oficina de universidad o centro de investigación— con el énfasis puesto en la interpretación y la escritura acerca de aquello observado en el "allí". Las notas de campo serán clave en esta instancia, y por lo tanto deben realizarse con máximo esmero y sentido crítico; en cuanto insumo de informaciones empíricas para el análisis intensificado típico del "estar aquí", ellas son decisivas a fin de facilitar una interpretación en la que no se pierda de vista "cómo se vinculan las palabras con el mundo, los textos con la experiencia, las obras con las vidas" (Geertz, 1988/1989: 145). Se trata al menos de evitar un relato ingenuo en el que lo recabado a través de las notas observacionales se tome como reflejo objetivo de los fenómenos estudiados.

CAPITULO 11

EL SONDEO

11.1. Introducción

El sondeo o encuesta ha sido definido como un método científico de recolección de datos, a través de la utilización de cuestionarios estandarizados, administrados por entrevistadores especialmente entrenados o distribuidos para su autoadministración a una muestra (Kuechler 1998).

Se trata de una técnica que permite recolectar datos sobre actitudes, creencias y opiniones de los individuos estudiados e indagar sobre múltiples temas, tales como pautas de consumo, hábitos, prejuicios predominantes e intenciones de voto en grandes poblaciones. Se caracteriza por su adecuación para relevar muchas propiedades referidas a muchos individuos. Y sus ámbitos de aplicación son diversos: académicos, políticos y comerciales.

Si bien los antecedentes pueden rastrearse en el siglo XIX (Marsh 1982), sus primeras aplicaciones como método moderno de recolección de datos, basado en la teoría estadística y las técnicas de muestreo, se sitúan en la década del treinta del siglo XX. En esos años, George Gallup aplicó la encuesta, tanto a estudios de mercado como de opinión pública.

La incorporación del sondeo a la investigación académica adquirió desarrollo recién durante los años cuarenta. Dentro del campo de la psicología social comenzó a utilizarse con mayor frecuencia en los estudios sobre la Segunda Guerra Mundial, cuando se llevaron a cabo importantes investigaciones entre las que se destaca la clásica obra dirigida por Samuel Stouffer (*El soldado americano*), basada en entrevistas a más de medio millón de soldados estadounidenses. En la Universidad de Columbia a partir de 1940, adquirió desarrollo y predominio la investigación social aplicada basada en encuestas en el campo de la sociología y los estudios de comunicación a través de los trabajos de Paul Lazarsfeld (1944, 1954). Durante esa década se crearon en Estados Unidos centros de investigación orientados a estudios cuantitativos basados en la técnica de encuesta, como el *National Opinion Research Center* (1941) y el *Survey Research Center* de Michigan (1946), y a partir de la década siguiente proliferaron las investigaciones académicas basadas en sondeos.

Su importante difusión en los Estados Unidos la convirtió en la técnica

más utilizada en ese país dentro de la investigación social empírica. Un estudio de los años ochenta, sobre el uso de encuestas en artículos publicados en revistas científicas, indica un aumento importante en el uso de la encuesta durante la década de 1960.

Tabla 11.1: *Porcentaje de artículos basados en encuestas en revistas de sociología y ciencia política (Estados Unidos)*

	1949-1950	1964-1965	1979-1980
tres revistas científicas líderes de sociología	24,1%	54,8%	55,8%
revistas de ciencia política	2,6%	19,4%	35,0%

Fuente: Presser 1985.

Según los datos de la tabla 11.1, los porcentajes de artículos basados en encuestas son crecientes, tanto en las revistas de sociología como en las de ciencia política, a lo largo del tiempo. Sin embargo, en las revistas de sociología analizadas el momento de mayor crecimiento es previo al mayor crecimiento en las de ciencia política. Asimismo, los trabajos publicados basados en encuestas constituyen un porcentaje mayor entre las revistas de sociología que entre las de ciencia política. Los datos muestran en alguna medida la tendencia creciente del uso de encuestas en el ámbito académico en ese país.

Debido a su éxito y aceptación, el método se extendió hacia otros países industrializados, como Inglaterra, Alemania, Francia y Japón, constituyendo la herramienta típica de los métodos cuantitativos en la medida que permite generar apoyo empírico para la inferencia estadística y estimar predicciones.

La base estadística de la encuesta convirtió a esta técnica en una de las más potentes para establecer ciertas predicciones basadas en datos empíricos, al demostrar que pequeñas muestras probabilísticas correctamente seleccionadas podían proporcionar algún grado de representatividad respecto de actitudes, opiniones y comportamientos de grandes poblaciones. En consecuencia, el principal atractivo de la encuesta reside en su potencial predictivo, al permitir el registro de alguna propiedad de una población estudiada, con un margen de error muestral calculable, a partir del análisis de un fragmento de ésta. Sin embargo, es importante destacar que, tratándose de estudios basados en probabilidades, los resultados que se obtienen a través de encuestas son estimaciones (aproximaciones) y no certezas. Aunque son muy pocos los estudios que permiten hacer estimaciones legítimas de los valores poblacionales, ya que esto sólo es posible cuando se trabaja con muestras probabilísticas (véase apartado 6.3). Por otro lado, también son pocos los escenarios donde es posible confrontar las predicciones resultantes de la encuesta con los parámetros poblacionales, pues esto exige relevar los estados en la propiedad bajo estudio en la población en su conjunto para luego compararla con el dato muestral. Obviamente, este procedimiento le quitaría sentido a la encuesta, ¿para qué estudiar una muestra

pudiendo estudiar toda la población? Un tipo de investigación que permite esta validación son los estudios preelectorales donde las estimaciones de las encuestas pueden ser confrontadas con el resultado de los comicios.

El sondeo, como toda técnica, es falible, tiene limitaciones, y en el proceso de su aplicación se acumulan errores (véase apartado 7.2). La minimización de éstos depende de la toma de decisiones adecuada y de controles permanentes en todas las etapas de su aplicación. Estos controles se orientan fundamentalmente a cuestiones relativas al ajuste y validez de los instrumentos, entre ellos: la adecuación de las preguntas, la coherencia del cuestionario, la validez de los indicadores, la significación de la muestra, la coherencia de los libros de códigos y la idoneidad de los encuestadores.

Existe, además, un conjunto de elementos que escapan al control técnico y que juegan un papel muy importante en la adecuación de los resultados finales de las encuestas. Se vinculan a condiciones relativas al talento y las aptitudes del investigador: la pericia para analizar los datos, la creatividad para elaborar factores de corrección y técnicas de proyección, la imaginación para conjeturar.

11.2. Supuestos

Pierre Bourdieu (1973) señala tres supuestos inverificables de las encuestas de opinión pública:

1) Todos los individuos tienen una opinión.
2) Todos los individuos se cuestionan sobre los temas que se les pregunta.
3) Todas las opiniones tienen la misma incidencia social.

Los dos primeros supuestos aluden a la influencia que la situación de encuesta puede tener sobre los encuestados, en el sentido que el interpelado se puede sentir compelido a responder y, a la vez, pensar que la expresión de su desconocimiento redundará en una evaluación negativa de su persona. Dados los límites de las técnicas para controlar la verdad de las respuestas obtenidas en las encuestas de opinión, en particular cuando se trata de preguntas cerradas, la cuestión de la veracidad de las respuestas/datos se vuelve un problema irresoluble. En consecuencia, lo que el individuo expresa es siempre considerado como la opinión de un sujeto que dice la verdad.

A fin de minimizar los riesgos de falta de veracidad de las respuestas, las encuestas tienden a realizarse en forma anónima y los investigadores garantizan la confidencialidad de los datos. En muchos países la confidencialidad de las respuestas de estudios oficiales basados en encuestas está protegida en el sistema legal. Generalmente el ámbito de aplicación de estas normativas y salvaguardas son los censos y los estudios realizados desde oficinas de estadística estatales. También se encuentran códigos de ética relativos a la confidencialidad en organizaciones de profesionales de encuestas.

Dentro de estas normativas, el tipo de recomendaciones y sugerencias más usuales se refieren a la protección de nombres, direcciones y teléfonos, o cual-

quier otro dato que permitiera la identificación del respondente, cualquiera sea el soporte de la información. También se aconseja, por un lado, que los datos se presenten en categorías suficientemente amplias de modo que ningún respondente pueda ser individualizado dentro de éstas y, por otro, la destrucción de los materiales identificatorios luego de volcarlos a las bases de datos.

El tercer supuesto de Bourdieu se funda en las investigaciones sobre grandes poblaciones, donde cada individuo es considerado equivalente a los demás y, en consecuencia, unos son sustituibles por otros. Esto permite trabajar matemáticamente las opiniones bajo el supuesto de la acumulatividad de las respuestas, pero, simultáneamente, se pierde la información relativa a la especificidad de cada sujeto (véase apartado 2.3). En este sentido, dentro de la encuesta tiene la misma importancia la opinión de un líder social que la de cualquier ciudadano. Algunos críticos de los sondeos (Blumer 1969; Habermas 1962), desde diferentes perspectivas, consideran que estos supuestos de la encuesta resultan en datos distorsionadores de la realidad. Ambos coinciden en que un método basado en la sumatoria de opiniones aisladas no permite explicar las relaciones en una sociedad donde los actores se encuentran integrados en diferentes colectivos sociales.

11.3. Limitaciones

Además de los supuestos señalados por Bourdieu, el sondeo tiene otras limitaciones. Algunas remiten a decisiones y estrategias investigativas previas a su aplicación: a) la validez de los indicadores y la confiabilidad de los instrumentos de registro, es decir el cuestionario mismo, y b) las limitaciones de las muestras utilizadas. La primera se intenta controlar, por un lado, testeando los indicadores antes de aceptarlos como válidos (véase apartado 9.4) y, por otro, sometiendo el cuestionario al control de una prueba piloto, mientras que la validez de la muestra depende de las técnicas de selección que se utilicen (véase capítulo 6).

Otras limitaciones operan durante el trabajo de campo, donde se generan la mayor parte de los errores que no se vinculan al diseño de los instrumentos sino a su aplicación. La mayoría de estos errores pueden minimizarse a través de un buen entrenamiento de los encuestadores y por medio de la supervisión y control de su trabajo.

Una limitación muy importante de las encuestas se produce cuando los encuestados se niegan a responder, debido a que una tasa de no respuesta muy alta puede vulnerar el cumplimiento de la muestra y sesgar los resultados afectando la calidad de los datos en totalidad (véase apartado 6.3). A fin de disminuir la tasa de no respuesta se pueden implementar estrategias de motivación y sensibilización en los encuestados, así como notificarlos con anterioridad sobre la visita del encuestador. Pero la posibilidad de llevar a cabo estas medidas depende en gran parte del tipo de población seleccionada y los recursos con que se cuente.

11.4. Tipos de encuesta según su abordaje

Denominamos abordaje a la técnica que se utiliza para llegar a los sujetos que constituyen el *target* del estudio. En relación con este criterio, podemos clasificar las encuestas en:

- con participación de encuestador
  - cara a cara o presenciales (domiciliarias y coincidentales)
  - telefónicas
- sin participación de encuestador (autoadministradas)
  - telefónicas automatizadas
  - correo tradicional
  - fax
  - correo electrónico/ *website*

Hasta los años setenta, el método utilizado en las encuestas era el presencial (cara a cara). A partir de entonces, comenzaron a implementarse las encuestas por correo y telefónicas (Dillman 1978), a las que actualmente se han sumado las encuestas por fax e Internet (*e-mail* o *websites*) (Dillman 2000).

Tabla 11.2: *Características de diferentes tipos de encuestas*

	PRESENCIALES	TELEFÓNICAS	CORREO	INTERNET
COSTO	ALTO	BAJO	MEDIO	MUY BAJO
encuestadores	sí	sí (no cuando son automatizadas)	no	no
entrenamiento de los encuestadores	alto	menor	no existe	no existe
posibilidades de supervisión	menores	mayores	no existe	no existe
rapidez	menor	mayor	depende de los respondentes	mucho mayor
cuestionario	complejo	simple	simple	simple
nº de preguntas	alto	bajo	medio	medio
distorsión en la formulación de las preguntas	mayor	menor (con supervisión)	ninguna	ninguna
intensidad del contacto personal	alto (comunic. visual y verbal)	bajo (comunic. verbal)	ninguno (com. escrita)	ninguno (com. escrita)
utilización de material extra	alta	nula	alta	alta
tasa de respuesta	depende de la población	depende de la población	depende de la población	depende de la población

11.4.1. ENCUESTAS CARA A CARA

La presencia del encuestador permite aclarar dudas, testear productos y mostrar piezas publicitarias u otros materiales. Además, en el caso de encuestas domiciliarias aumentan las posibilidades de utilizar un cuestionario más largo, dado que las condiciones de la entrevista son más confortables para el encuestado que cuando es abordado en la vía pública. La principal ventaja de la encuesta domiciliaria es la posibilidad de introducir un elemento de azar en el diseño de la muestra, a través del sorteo de unidades espaciales de un mapa de la zona a encuestar; su principal desventaja es que resulta más costosa.

Una función muy importante del encuestador se encuentra en el caso de muestras con cuotas, es decir, cuando se construye una muestra que reproduce las proporciones de alguna/s variable/s de la población (por ejemplo, sexo y edad) donde los encuestadores son los encargados de aplicar las cuotas que se hayan establecido.

El método coincidental (abordaje en la vía pública o lugares públicos) resulta eficiente cuando se estudia un segmento poblacional que puede ser detectado por su concentración espacial (por ejemplo, seguidores de un club de football encuestados en el estadio o consumidores encuestados en el supermercado). También, algunos estudios en EE.UU. han demostrado la eficiencia de esta estrategia de abordaje para llegar a segmentos de difícil acceso, como el estudio de jóvenes afroamericanos entrevistados en puntos de reunión en la calle. En comparación con los estudios telefónicos, estos diseños resultaron en mejores distribuciones en edad y sexo (Miller y Wilder 1997). En general, las encuestas cara a cara resultan más útiles para segmentos de bajos ingresos, por su falta de teléfono y porque producen mayores tasas de respuesta.

Es importante tener en cuenta que el lugar donde se aborda al entrevistado puede tener influencia en la calidad de los datos. En un estudio en aeropuertos donde las unidades de análisis eran pasajeros de viajes aéreos, se hizo un análisis comparativo abordando a los sujetos en la puerta (*gate*) y a bordo del avión (Denstadli 2000). Los resultados indicaron que la tasa de no respuesta era mayor en el *gate,* probablemente debido a un mayor estado de intranquilidad del encuestado, mientras que la aplicación del cuestionario a bordo presentaba la ventaja de una mayor disponibilidad de tiempo y una mejor disposición de los pasajeros.

11.4.2. ENCUESTAS TELEFÓNICAS

Las encuestas telefónicas se han vuelto cada vez más usuales, particularmente en zonas urbanas donde altos porcentajes de la población tienen teléfono. Para su realización se necesita tecnología informática *ad hoc* (CATI, *Computer Assisted Telephone Interviewing*) que permite el marcado de números al azar y el procesamiento de los datos inmediatamente de completado el trabajo de campo.

Con el desarrollo del IVR (*Interactive Voice Responder*) es posible diseñar encuestas telefónicas automatizadas (sin encuestador). Se trata de un sistema

informático capaz de grabar y reproducir archivos de audio y reconocer respuestas a través de los tonos del teléfono.

La principal ventaja de la encuesta telefónica es el menor costo y la rapidez. Sus desventajas son: la imposibilidad de alcanzar a la población sin teléfono (que tiende a ser la de menores ingresos), la restricción a un cuestionario corto y la imposibilidad de utilizar material visual.

11.4.3. ENCUESTAS AUTOADMINISTRADAS (SIN ENCUESTADOR)

Se denominan "autoadministradas" aquellas en las que no participan encuestadores sino que se provee el cuestionario al encuestado para que él complete los datos que se solicitan.

Dentro de la población a estudiar se selecciona una muestra potencial y se la invita a participar en el estudio haciéndole llegar los cuestionarios (a través del correo postal, correo electrónico o dejándolos en alguna institución, empresa u organismo al que pertenezcan los encuestados potenciales). Una vez que el cuestionario llega a los sujetos, su respuesta y su devolución dependen de su voluntad. De manera que la muestra real del estudio resulta de la autoselección de los entrevistados. Esto generalmente se traduce en una baja tasa de respuesta (aunque ésta depende de las características de la población y su interés en la temática). Habitualmente, con la finalidad de aumentar el número de respuestas, se utilizan técnicas que pueden consistir desde premios hasta compartir los resultados del estudio. Generalmente, la tasa de respuesta tiende a ser mayor en poblaciones con alta educación e interesadas en la temática de que se trate.

Es importante estimar la cantidad de respuestas de antemano a fin de calcular el número de cuestionarios que deben enviarse para completar la muestra. Cuando la muestra potencial inicial es muy heterogénea, la autoselección puede producir sesgos, es decir, que la muestra obtenida no refleje esa heterogeneidad.

En las encuestas autoadministradas el cuestionario adquiere ciertas características. La ausencia de encuestador exige la redacción de un cuestionario de poca extensión y con consignas muy claras que puedan ser seguidas con precisión por el encuestado al completarlo. Entre las posibilidades que ofrece este tipo de encuestas se cuenta el envío de imágenes, pero es necesario considerar que esto aumenta los costos.

Cuando las encuestas autoadministradas son difundidas por correo electrónico o en sitios de Internet resultan más rápidas y económicas; sin embargo, la utilización de tecnología que exige el acceso de los encuestados a Internet limita mucho los *targets* y, en consecuencia, la selección de la muestra. A fin de conformar ésta, es necesario contar con un listado de direcciones electrónicas de individuos, empresas o instituciones que integren la población en estudio. Es importante destacar, junto a estas limitaciones, las potencialidades de esta estrategia en la medida que Internet permite llegar a gran parte de los países, aunque todavía existan muchos problemas para la implementación de encuestas a través de esta vía en los países de menor desarrollo.

Desde fines de los años ochenta se llevaron a cabo estudios comparativos de la eficiencia de las encuestas por correo tradicional y por correo electrónico. Según un estudio de Lee Sproull (1986), basado en cuestionarios por *e-mail* y en papel a dos muestras de treinta individuos cada una, en una empresa, las respuestas por *e-mail* llegaron en la mitad del tiempo que las respuestas en papel, pero éstas eran más completas. Duane Bachmann, John Elfrink y Gary Vazzana (1996) también compararon encuestas por *e-mail* y por correo. Sus conclusiones indican que el *e-mail* tiene buena llegada a determinados grupos sociales, particularmente profesionales, pero no es posible llegar a muchos otros sectores sociales por esta vía. En general los respondentes vía *e-mail* son más jóvenes y mayores consumidores.

11.5. Estudios comparados (*cross-national, cross-cultural studies*)

Este diseño consiste en aplicar encuestas en diferentes países o culturas y replicarlas sistemáticamente a intervalos regulares. Constituyen bancos de datos importantes por su alcance y sistematicidad. Entre las más conocidas se encuentra el *International Social Survey Program* (ISSP) y el Eurobarómetro. En América latina, desde la última década, se aplica el Latinobarómetro.

El Eurobarómetro incluye los países de la Unión Europea, el ISSP es más diverso, incluye países como Bangladesh, Chile y Filipinas junto con Australia, Inglaterra, EE.UU. y Alemania, mientras que el ámbito del Latinobarómetro está conformado por dieciocho países de América latina.

El principal problema que presentan estos estudios es la comparabilidad de los datos, cuando incluyen realidades culturales, sociales, económicas o políticas muy heterogéneas. Por esta razón Manfred Kuechler (1998) recomienda que se incluyan componentes cualitativos en los diseños. En general se sugiere:

- realizar ajustes en términos culturales;
- buscar similitud en los países seleccionados respecto de las características institucionales que afectan los temas centrales del estudio;
- la presencia de, al menos, un investigador de cada país con manejo del idioma y conocimiento de las costumbres en el equipo de investigación, en particular para el diseño del cuestionario;
- analizar primero cada realidad nacional y luego hacer el análisis comparado;
- recolectar los datos en diferentes momentos, para neutralizar el efecto de eventos nacionales.

11.6. Estudios longitudinales

La estrategia metodológica del PANEL reside en la utilización de la muestra. El cuestionario se aplica a una misma muestra en diferentes ocasiones a lo largo del tiempo, de modo que resulta particularmente útil para comprobar modificaciones en las actitudes y conductas como efecto de la participación de un estímulo o factor determinado. Uno de los principales problemas del *panel* se vincula a la mortalidad de la muestra, en la medida que puede perderse el contacto con algunos individuos de un registro a otro.

Lazarsfeld, Berelson y Gaudet (1944) aplicaron en forma innovadora una estrategia de panel en el estudio de una campaña electoral, que les permitió detectar las fluctuaciones en las opiniones de los electores y los efectos de la propaganda en sus actitudes políticas.

11.7. El cuestionario

El cuestionario es el instrumento de recolección de datos propiamente dicho utilizado en la técnica de encuestas. Es importante tener en cuenta que su elaboración no consiste en diseñar un simple listado de preguntas; por el contrario, su estructura tiene una lógica que es conveniente respetar. La idea general es que las primeras baterías de preguntas no resulten demasiado intrusivas ni requieran respuestas comprometidas, de manera de generar confianza y un clima amigable para el respondente. Los temas que exigen mayor compromiso o apuntan a núcleos de mayor intimidad deben ubicarse hacia el centro de la entrevista, para terminar con preguntas fáciles que den lugar a la despedida del encuestador en un clima de amabilidad.

Por otro lado, la duración de la entrevista no debe ser excesiva para que el encuestado pueda mantener la atención en los temas que se le proponen. A fin de no saltar de un tema a otro, cosa que puede provocar confusión, es conveniente dividir los cuestionarios en dimensiones temáticas, buscando agotar un tema antes de presentar otro.

Según su grado de estructuración, los cuestionarios se clasifican en estructurados, semiestructurados y no estructurados. En los primeros tienden a predominar las preguntas que presentan un menú cerrado de respuestas precodificadas (denominadas PREGUNTAS CERRADAS). Éstas otorgan muy poca libertad al encuestado para desarrollar sus respuestas, quien fundamentalmente elige entre un listado de alternativas que se le presenta.

El cuestionario estructurado responde a los supuestos del método hipotético deductivo, en tanto la investigación se concibe como un proceso para contrastar hipótesis preestablecidas y no para explorar la realidad a través de la inducción. Por otro lado, presentar a todos los encuestados las mismas preguntas con los mismos menúes de respuestas se orienta a asegurar, en alguna medida, la comparabilidad de éstas, ya que los estímulos han sido los mismos para todos los integrantes de la muestra. Obviamente, este supuesto tiene la debilidad de que si bien la pregunta puede ser idéntica —suponiendo que el encuestador la formule siempre en los mismos términos del cuestionario— la situación de entrevista es siempre diversa.

El cuestionario semiestructurado es similar al anterior, pero incorpora algunas PREGUNTAS ABIERTAS, donde el encuestado tiene libertad para elaborar su propia respuesta en sus propias palabras. Estas preguntas cumplen una fun-

ción diferente, responden a una actitud exploratoria del investigador. Es decir, tienden a buscar información desconocida a través del camino de la indagación. Los cuestionarios semiestructurados son los más utilizados e incorporan, a través de las preguntas abiertas, un elemento de naturaleza no estándar (también llamada cualitativa) dentro de una técnica estándar (también llamada cuantitativa).[1] En la etapa del procesamiento de la información estas preguntas deben ser "cerradas", es decir, a través de un proceso interpretativo se elabora una clasificación de las respuestas a cada pregunta abierta y se asigna un código numérico a cada tipo de respuesta.

En el cuestionario no estructurado la mayoría de las preguntas son abiertas y los entrevistados tienen libertad para elaborar sus respuestas en forma y contenido. Este tipo de cuestionario es utilizado fundamentalmente en las entrevistas en profundidad individuales y grupales, donde adquiere el nombre de pauta de entrevista (véanse los capítulos 12 y 13).

11.8. Formulación de las preguntas

La primera clasificación de las preguntas de un cuestionario las divide en abiertas y cerradas. En el diseño de las preguntas cerradas se pueden presentar las respuestas posibles (que se someterán a la elección del encuestado) en diversas formas. Cuanto mayor es el número de alternativas que se presentan mayores posibilidades de elección tiene el respondente, mientras que cuanto menor es este número más se lo obliga a encasillar su respuesta dentro de un abanico poco diverso y, en consecuencia, en categorías más amplias.

El diseño de las respuestas para cada pregunta cerrada estará condicionado por el grado de conocimiento del investigador sobre la materia, el nivel de heterogeneidad de la muestra y la amplitud en la dispersión de las respuestas que se esté dispuesto a tolerar.

La pregunta que presenta menor alternativa de respuesta es la DICOTÓMICA: las respuestas posibles son sólo dos, del tipo sí/no o acuerdo/ desacuerdo. Este tipo de pregunta tiende a dividir la muestra en dos grupos en oposición, relativos a una temática (los que creen y los que no creen, los que aceptan y los que no aceptan, los que pertenecen y los que no pertenecen).

Cuando la pregunta presenta más de dos alternativas de respuesta entre las cuales el respondente elegirá una sola, se denomina de ALTERNATIVAS MÚLTIPLES. Como se dijo en el capítulo 8, las alternativas múltiples pueden ser categoriales (por ejemplo, ante una pregunta sobre principales problemas, ofrecer como respuestas: económicos, políticos o sociales) u ordinales, que generalmente asumen formas escalares. Las escalas, a su vez, pueden ser verbales (muy buena, buena, regular, mala, muy mala) o numéricas (solicitar al respondente que posicione su respuesta en una escala 1-10). Cuando en una pregunta de alternativas múltiples se permite elegir más de una respuesta (es-

[1] Sobre los términos 'cualitativo/cuantitativo' y 'estándar/ no estándar' véase el capítulo 2.

to no ocurre en el caso de las escalas), adopta el nombre de pregunta de RESPUESTAS MÚLTIPLES.

Generalmente, cuando el conjunto de respuestas posibles es finito, conocido y no es demasiado grande (por ejemplo los partidos políticos que presentaron candidatos en la última elección presidencial) se puede formular al entrevistado una pregunta abierta (¿Qué partido votó en la última elección presidencial?) sin presentarle el menú de respuestas, y el encuestador registra la respuesta en el listado de partidos contenido en el cuestionario. De estas preguntas se dice que se formulan en forma ESPONTÁNEA. En los casos en que un menú de respuestas es sometido a consideración del encuestado, se trata de preguntas formuladas en forma GUIADA.

Cuando una pregunta cerrada tiene un número grande de categorías de respuesta que se vuelven difíciles de recordar, se puede utilizar una 'tarjeta' —instrumento que ayuda a guiar al encuestado— donde figuran las alternativas de respuesta que aquél lee antes de decidir su elección.

El conjunto de respuestas de una pregunta cerrada constituye una clasificación y debe cumplir con los requisitos de ésta: todas las respuestas deben ser mutuamente excluyentes (no superponerse en ningún valor) y conjuntamente exhaustivas (contener en conjunto todas las respuestas posibles) (véase apartado 8.1). A fin de cumplir con esta última condición, toda pregunta llevará una respuesta residual ('otra') y una categoría para registrar la falta de respuesta ('ns/nc', no sabe/no contesta).

Otra estrategia para recolectar información a través de un cuestionario es la formulación de frases. Ésta consiste en redactar un conjunto de frases que expresan diferentes posiciones respecto de alguna temática y solicitarle al encuestado que manifieste su opinión sobre ellas (generalmente la respuesta adopta la forma de una escala Likert, véase el apartado 8.5). Esta técnica tiende a utilizarse cuando se indagan temas controversiales o conflictivos, su ventaja es que las frases son presentadas como dichas por otras personas y generan en el respondente una sensación de menor compromiso al considerarse formando parte de algún grupo.

Cuando se quiere indagar acerca de preferencias, se puede presentar al encuestado un listado a fin de que lo ordene o jerarquice sus elementos según algún criterio (por ejemplo, ordenar una lista de dirigentes políticos según su capacidad para resolver determinado problema). Otro tipo de preguntas llamadas FILTRO O DE PASE son aquellas cuya respuesta reorienta a diferentes segmentos de la muestra en las preguntas del cuestionario. A partir de éstas, el orden de presentación de las preguntas se altera para algún segmento de la muestra porque para el mismo algunas preguntas se vuelven irrelevantes (por ejemplo, la indagación sobre afiliación sindical puede dividir a la muestra en afiliados y no afiliados, y la batería de preguntas inmediatamente siguiente puede ser relevante para los primeros e irrelevante para los segundos). Algunas veces se ubican preguntas filtro al inicio del cuestionario a fin de decidir la inclusión o exclusión de sujetos en la muestra (por ejemplo, si se trata de una muestra de ciudadanos del distrito A, una pregunta filtro podría ser ¿Usted en qué distrito vota? Sólo se aplicará el cuestionario a quienes respondan: "En el distrito A").

Al final del cuestionario se ubican las llamadas PREGUNTAS DE CLASIFICACIÓN, orientadas a recolectar datos sociodemográficos de los encuestados, como edad, educación, estado civil y ocupación. Estas preguntas permiten caracterizar la muestra y compararla, generalmente con parámetros poblacionales censales, a fin de evaluar su grado de aproximación en estas variables. Como su nombre lo indica, se utilizan para clasificar la muestra, segmentándola según estas variables, a fin de estimar sus efectos en las opiniones dado su carácter más estructural o más permanente.

Tanto el orden de las preguntas en el cuestionario como el orden que adopten las respuestas al interior de cada pregunta cerrada pueden influir en los resultados. A fin de controlar estos efectos se recomienda realizar una prueba piloto del cuestionario y alterar la presentación de las respuestas entre los encuestados, en particular cuando tienen muchas alternativas. De este modo, se minimizan los efectos de las respuestas sistemáticas, es decir, la tendencia de algunos respondentes a elegir siempre una misma alternativa (generalmente la primera o la última que se les presenta).

A la hora de diseñar un cuestionario es importante tener en cuenta que la influencia de las preguntas sobre las respuestas está condicionada en gran parte por el lenguaje. Habitualmente, se da por supuesto que el interpelado reconocerá el sentido de la pregunta tal como fue formulado —"supuesto de la interpretabilidad" (Clark y Schober, 1991)— pero esto no es necesariamente así. El lenguaje utilizado conlleva una intención y una perspectiva que no siempre son compartidas por los participantes en el sondeo. El interpelado reconstruye el sentido de la pregunta desde su propia perspectiva cultural, y su respuesta responde a esta pregunta reconstruida. Este proceso, propio de toda actividad conversacional, afecta la validez de las respuestas de la encuesta y su comparabilidad. El problema se agrava cuando no existen campos compartidos de sentido ni usos semejantes del lenguaje entre quien redacta las preguntas y los respondentes o —en el caso de una muestra culturalmente heterogénea— entre los respondentes entre sí.

Muchas palabras utilizadas frecuentemente en las encuestas tienen una carga importante de ambigüedad e imprecisión (por ejemplo, mayor, menor, bueno, muy bueno, malo, suficiente) y pueden adquirir sentidos diferentes para los encuestados. También puede ocurrir que se utilice un lenguaje técnico en las preguntas que muchos entrevistados no comprendan. Se vuelve importante entonces, en el diseño del cuestionario, minimizar la utilización de palabras y expresiones vagas o ambiguas que agraven el distanciamiento que pueda existir entre la perspectiva de quien redacta el cuestionario y los encuestados.

CAPÍTULO 12

LA ENTREVISTA EN PROFUNDIDAD

12.1. Introducción

El término 'entrevista' en el sentido actual —al menos en cuanto entrevista periodística— pareciera haber sido empleado por primera vez en la lengua inglesa (*interview*) en un artículo del semanario neoyorquino *The Nation*, de principios de 1869. Sin embargo, en cuanto verbo cuyo significado alude a "tener un encuentro personal", el término existe en dicho idioma desde al menos 1548.[1] Este significado remite a su origen etimológico, ya que, aunque derivado más inmediatamente de la palabra francesa *entrevue*, sus raíces pueden rastrearse hasta los términos latinos *inter* y *videre*, literalmente "entre ver", y de ahí "verse mutuamente", "tener un encuentro cara a cara".

En las ciencias sociales la entrevista se refiere a una forma especial de encuentro: una conversación a la que se recurre con el fin de recolectar determinado tipo de informaciones en el marco de una investigación. Tal como reportan Fideli y Marradi (1996), se trata de la técnica más utilizada en estas disciplinas, con estimaciones que indican que cerca del 90% de los estudios empíricos se valen en algún modo de ellas.

Un simple ejercicio nos ha demostrado que estas estimaciones no son para nada exageradas: analizando cuantitativamente los *abstracts* de los artículos indizados en *Sociological Abstracts*,[2] en los que habitualmente se reportan los medios (métodos y técnicas) empleados para llevar adelante la investigación, hemos detectado que en más del 43% de casi 117 mil trabajos se usa el término 'entrevista' en sentido estricto. Si aceptamos que también las encuestas/sondeos (que representan otro 43%) involucran una forma especial de entrevista

[1] Fuente: *Online Etimology Dictionary* de Harper, 2001, <www.etymonline.com> (fecha de consulta: 14/9/2006).

[2] Se trata de una base de datos que compila la literatura científica internacional de las ciencias sociales publicada en más de 4.300 revistas de todo el mundo (además de libros y actas de congresos) y editada por *Cambridge Scientific Abstracts*. Su cobertura temporal se expande desde los años cincuenta hasta la actualidad. Los datos consignados en este capítulo se basan en una consulta realizada el día 7 de diciembre de 2006.

—como se verá más adelante, en el apartado 12.2—, tenemos que en más del 86% de los estudios empíricos indizados los investigadores se han valido de entrevistas. Si además se reconoce que la realización de observaciones participantes (3,9%), habitualmente desarrolladas en el marco de estudios etnográficos, igualmente conlleva algún tipo de entrevista, llegamos a superar el 90% del total. Finalmente, dado que en los experimentos (9,7%) también se puede recurrir a dispositivos genéricamente clasificables como entrevistas para la recolección de información, concluimos que las estimaciones citadas por Fideli y Marradi (1996) en torno del peso relativo de la entrevista en la investigación social empírica son más que razonables.

Sin embargo, cabe hacer notar que nos hemos referido a "algún uso" de entrevistas y no al "uso exclusivo" de ellas. En este sentido, resulta oportuno señalar que, especialmente en la tradición antropológica, la entrevista ha sido vista en general como un complemento de la observación de campo (práctica a la que se le concede un papel preponderante), mientras que en la sociología y otras ciencias sociales —aun sin negar la posibilidad de complementar/articular el uso de entrevistas con otras técnicas de recolección— se les ha dado un lugar central y valor científico por sí mismas.

En tanto forma de conversación, la vida cotidiana y los procesos de socialización —como plantean Valles (1997) y Alonso (1998), entre otros— pueden constituirse en la base para la caracterización preliminar y el aprendizaje de la entrevista. En efecto, en la vida cotidiana estamos recurrentemente involucrados en prácticas conversacionales que implican el intercambio de información; las competencias para desempeñarnos en ellas, que se adquieren a través de la socialización, empezando por el lenguaje y las capacidades del habla, y siguiendo por el conjunto de saberes tácitos que hacen posibles tales interacciones verbales, son cruciales también en la práctica de la entrevista de investigación social.

Por otra parte, cabe señalar que además de las conversaciones informales a las que acabamos de referirnos, existen también en la vida cotidiana intercambios más formalizados que suelen definirse como entrevistas profesionales: la consulta médica, la entrevista psicoterapéutica, la de selección de personal, la periodística, etcétera. La entrevista de investigación social podría concebirse como una forma especial de conversación profesional. Se diferencia de la entrevista psicoterapéutica, por ejemplo, en que su objetivo primario es conocer y no, en cambio, modificar las actitudes o conductas del entrevistado. Este fin eminentemente cognitivo la emparienta con otros tipos de entrevista —tómese por caso la periodística—, pero a diferencia de ella se desarrolla en el marco de una investigación científica o académica (Fideli y Marradi 1996).

En definitiva, siguiendo a Alonso (1998), podemos definir la ENTREVISTA EN PROFUNDIDAD como una forma especial de conversación entre dos personas (aunque como se verá más adelante también existen formas de entrevista grupal), dirigida y registrada por el investigador con el propósito de favorecer la producción de un discurso conversacional continuo y con cierta línea argumental por parte del entrevistado, acerca de un tema de interés definido en el marco de la investigación.

12.2. Tipos de entrevista

En las ciencias sociales existen distintos tipos de entrevista y hay a su vez diferentes fundamentos para clasificarlas. Fideli y Marradi (1996) proponen una tipología basada en dos criterios: a) la presencia (o no) de un contacto visual directo entre entrevistador y entrevistado, y b) el grado de libertad concedido a los actores en la situación de entrevista, ya sea al preguntar o al responder. Este último criterio es probablemente el más utilizado en la clasificación de la entrevista, dando lugar a un *continuum* de grados de libertad posibles en función del nivel de estructuración o estandarización que adquiera la guía de preguntas —y/o de las posibles respuestas predeterminadas— a ser empleada en la entrevista.

Burgess (1984) afirma que el grado de menor estructuración se da en el marco de las conversaciones espontáneas que surgen como complemento de la observación de campo. En el otro extremo se ubican las entrevistas estructuradas típicas de las encuestas o sondeos, en las que se aplica de modo uniforme un cuestionario estandarizado que contiene todas las preguntas a realizar —y en un orden dado— y todas (o casi todas) las opciones posibles de respuesta (véase capítulo 11).

Si bien la estructuración —como se acaba de decir— se puede pensar a la manera de un *continuum* entre los polos recién mencionados, a los efectos de simplificar la cuestión presentamos una tipología en la que el grado de espontaneidad de la interacción verbal (directamente relacionado con el nivel de estructuración previa de preguntas y respuestas) se limita a tres puntos —bajo, medio y alto— dando lugar a formas estructuradas, semiestructuradas y no estructuradas de entrevista. Por otro lado, el criterio relativo al tipo de contacto entre entrevistador y entrevistado, que en el caso de Fideli y Marradi (1996) se limita a la relación personal cara a cara y al contacto telefónico, incluye también en nuestra propuesta la interacción virtualizada.[3]

Tabla 12.1: *Tipología de las entrevistas*

	GRADO DE ESPONTANEIDAD DE LA INTERACCIÓN VERBAL		
TIPO DE CONTACTO	BAJO	MEDIO	ALTO
personal (cara a cara)	entrevista personal estructurada	entrevista personal semiestructurada	entrevista personal no estructurada
telefónico	entrevista telefónica estructurada	entrevista telefónica semiestructurada	entrevista telefónica no estructurada
virtual	entrevista virtual estructurada	entrevista virtual semiestructurada	entrevista virtual no estructurada

[3] Es de hacer notar, sin embargo, que esta última podría adquirir diversos formatos, acercándose a una encuesta con cuestionario autoadministrado (si el uso de la tecnología informática se limita al envío por correo electrónico de tal instrumento); a una entrevista telefónica (si la tecnología se utiliza para comunicarse en tiempo real, pero sin contacto visual), y a una entrevista cara a cara (si se recurre a tecnologías multimedia que involucren audio y video). Por otra parte, también con las nuevas tecnologías de telecomunicaciones existen dispositivos telefónicos que permiten el contacto visual.

En la tabla se han resaltado en gris las formas de entrevista que genéricamente podrían considerarse "en profundidad" y que constituyen el objeto del presente capítulo. Dado que las versiones telefónicas y virtuales se encuentran menos difundidas, nos concentraremos exclusivamente en las entrevistas cara a cara. Las formas estructuradas, por su parte, ya han sido tratadas en el capítulo 11. Por otra parte, cabe destacar que hasta ahora hemos dado por supuesto un tipo de relación diádica entre un entrevistador y un entrevistado; sin embargo, existen también formas de entrevista grupal. La más difundida —el *focus group*— será tratada en el capítulo 13.

Antes de pasar a la caracterización detallada del tipo de entrevistas que aquí nos ocupa resulta pertinente señalar que no existe aún en la metodología de las ciencias sociales una terminología unificada para designarlas. A lo largo de los años se han utilizado distintas expresiones para referirse a modalidades que podrían encuadrarse genéricamente bajo el rótulo de entrevistas en profundidad. Hemos decidido recurrir justamente a este término para englobar las distintas formas específicas a las que aludimos, aun reconociendo la existencia de sutiles diferencias entre ellas. Especialmente en el caso de las menos estructuradas se habla con frecuencia de entrevista abierta; no directiva (Rogers 1942); no estandarizada (Denzin 1970); intensiva (Brenner *et al.* 1985); cualitativa (Valles 1997); hermenéutica (Monsteperelli 1998). En el caso de las semiestructuradas, Combessie (1996) recurre a la expresión "entrevista centrada". Merton y Kendall (1946), por su parte, propusieron la noción de entrevista focalizada: un tipo semiestructurado que se caracteriza por el hecho de que los entrevistados han recibido un estímulo específico (han visto una película, leído un libro, etcétera) o han participado de una situación social cuya experiencia subjetiva es objeto de la entrevista. Marradi (2005) ha propuesto una forma especial de entrevista semiestructurada a partir de un estímulo particular; se trata de historias (*storie*) en las que se narran episodios de la vida cotidiana y que el entrevistador cuenta al entrevistado para luego explorar sus reacciones y reconstruir sus preferencias valorativas.

12.3. Caracterización de la entrevista en profundidad

Rosenblum (1987) afirma que la entrevista en profundidad se caracteriza por su estructura paradójica: en un intercambio explícitamente instrumental y muy circunscrito en el tiempo —entre personas que son relativamente extrañas— exige a la vez intimidad e impersonalidad, profesionalismo en un marco de sociabilidad.

Se trata fundamentalmente —como señala Alonso (1998: 67-68)— de "un *proceso comunicativo* por el cual el investigador extrae una información de una persona". Pero no cualquier tipo de información, sino aquella que se halla contenida en la biografía del entrevistado, es decir, aquella que se refiere "al conjunto de representaciones asociadas a acontecimientos vividos por [él]". En este sentido, la información que interesa al investigador ha sido experimentada e interpretada por el entrevistado; ésta forma parte de su mundo de la vida

—antes tácito, dado por descontado— y que ahora pasa a ocupar el centro de la reflexión, siendo problematizado y narrado (Montesperelli 1998: 73).

Guber (1991/2005: 236) destaca que "la entrevista es un proceso en el que se pone en juego una relación social", y ésta involucra a dos actores: el entrevistador y el entrevistado. Dado que tal relación suele ser en muchos sentidos asimétrica, el entrevistador debe ser capaz de reflexionar sobre su rol, sus elecciones y —de un modo más amplio— sobre la dirección y el sentido de su investigación. Para Montesperelli (1998), debe tener una persistente actitud de apertura, y estar dispuesto a encontrar aquello que no esperaba, moviéndose constantemente entre observación y conceptualización. El entrevistador tiene un rol estratégico y "mayéutico": no direcciona las respuestas; asume en cambio que el entrevistado es el verdadero experto en el tema objeto de la entrevista y que es capaz de explicitar su propio conocimiento, limitándose a "ayudarlo" —en un proceso dialéctico— a que lo verbalice, invitándolo a la conversación. Pero esto no es una cuestión sencilla: el entrevistador debe ser consciente de los problemas que implica la conversación y debe tender a comunicar con naturalidad y sensibilidad; su función primordial es la de escuchar.

Este tipo de entrevista se caracteriza entonces por el alto grado de subjetividad, hecho que para muchos constituye simultáneamente su principal rasgo y mayor limitación. Alonso (1998) habla incluso de "hipersubjetividad": el entrevistado es un *self* que relata historias mediadas por su memoria e interpretación personal, y en este sentido la información que provee no debe apreciarse en términos de veracidad-falsedad, sino como el producto de un individuo en sociedad cuyos relatos deben ser contextualizados y contrastados. Como indica Guber (1991/2005: 237), "un término, un discurso o una acción no son lo que son *per se*, sino en relación con la situación en que se enuncian o aplican y con su contexto discursivo y material". La entrevista, por lo tanto, sólo se puede juzgar "por la riqueza heurística de las producciones discursivas obtenidas en ella" (Alonso 1998: 77), las que, por otra parte, no son en sí mismas correctas ni incorrectas (Fideli y Marradi 1996).

Según la perspectiva de Alonso (1998), en la entrevista en profundidad no se expresa simplemente una sucesión de acontecimientos vividos sino la verbalización de una apropiación individual de la vida colectiva. Es por esto que no se trata de un registro fedatario de hechos o datos; es en cambio el arte del vínculo: un juego de estrategias comunicativas a partir del cual se registra un "decir sobre el hacer".

La relevancia o irrelevancia del resultado de este juego comunicativo depende en gran medida de las capacidades, la perspicacia y la personalidad del entrevistador (Statera 1982). No obstante, la viabilidad del juego se basa en el establecimiento de un contrato comunicativo que involucra un conjunto de saberes explícitos e implícitos —compartidos por entrevistador y entrevistado— que harán posible el funcionamiento de la entrevista. Entre los saberes explícitos se destacan aquellos relativos al tema objeto de la conversación y a los fines de la investigación en general: ¿quién, para qué y por qué realiza la entrevista? En este sentido, es importante destacar que ella debe realizarse a partir de un consentimiento informado de parte del entrevistado, y a su vez esto exige ex-

plicitar ante él las cuestiones que acabamos de referir. Los saberes implícitos, por su parte, aluden a los códigos lingüísticos, culturales y reglas sociales que hacen posible la comunicación interpersonal en un contexto dado; en definitiva, cabe recordar que la entrevista se funda en la interacción verbal, y requiere por lo tanto una apertura a la comunicación tanto como la aceptación de sus reglas (Alonso 1998).

12.4. Usos de la entrevista en las ciencias sociales. Ventajas y limitaciones

La entrevista se utiliza en general cuando se busca acceder a la perspectiva de los actores, para conocer cómo ellos interpretan sus experiencias en sus propios términos. En este sentido, Montesperelli (1998) resalta sus virtudes siempre que se trate de explorar el mundo de la vida cotidiana. Alonso (1998), por su parte, señala que la entrevista en profundidad es especialmente aplicable cuando se busca reconstruir acciones pasadas; estudiar representaciones sociales personalizadas; analizar las relaciones entre el contenido psicológico personal y la conducta social, o explorar campos semánticos, discursos arquetípicos de grupos y colectivos.

Valles (1997: 196-198) ha sistematizado de modo didáctico las ventajas y limitaciones del uso de la entrevista en profundidad en la investigación social. Con respecto a las primeras destaca que permite obtener de manera flexible una información rica y profunda, en las propias palabras de los actores. Asimismo, señala que proporciona la oportunidad de clarificación en un marco más dinámico y espontáneo que el de las entrevistas estructuradas de las encuestas. Si se piensa en las fases iniciales de un estudio, resultan significativas sus ventajas para realizar los primeros acercamientos al tema, y en el caso de las fases finales, se destaca su uso a fin de enriquecer los resultados de indagaciones cuantitativas o cualitativas, ya sea a través del contrapunto o de la comprensión más profunda de éstos. Comparada con la observación participante, Valles señala la capacidad de la entrevista para permitir el acceso a cierto tipo de informaciones difíciles de conocer sin la mediación del entrevistador. En este mismo sentido debería considerarse la posibilidad de conocer, a través del relato de los actores, situaciones no directamente observables —entre ellas aquellas del pasado. En tiempos recientes, reconociendo justamente esta cualidad de las entrevistas, se ha difundido notablemente su uso para el estudio del pasado próximo y las formas de apropiación individual y colectiva (memoria) de él. Finalmente, si se la contrasta con las entrevistas grupales, la entrevista en profundidad tiene la ventaja de la mayor intimidad.

Entre las limitaciones, Valles (1997) señala, con respecto a la encuesta, una desventaja relativa en términos de tiempo. A esto habría que agregarle otros elementos: menor capacidad para captar fenómenos con gran dispersión territorial y/o tipológica, menor capacidad para generalizar sus resultados, mayor complejidad en el registro, procesamiento y análisis de éstos. Al igual que otras técnicas basadas en la interacción comunicativa, la entrevista presenta problemas potenciales de reactividad, fiabilidad y validez. Comparada con la observación participante, se considera una restricción la imposibilidad de observar los fenómenos en su ambiente natural, limitación que Cicourel (1964) calificó como carencia de validez ecológica.

En definitiva, y a pesar de su utilidad para acceder —potencialmente— al universo de las significaciones de los actores: sistemas de representaciones, creencias, normas, valores, nociones, etcétera (Guber 1991), hay que reconocer que la efectiva realización de las ventajas de la entrevista, así como la minimización de sus debilidades, queda indefectiblemente condicionada por la capacidad de empatía del entrevistador y su habilidad para crear un clima que favorezca la comunicación. En esto influyen, tal como se dijo oportunamente en el caso de la observación participante (véase capítulo 10), las características personales del entrevistador; género, edad, clase social, etnia, etcétera, son cuestiones sensibles que pueden potenciar tanto las ventajas como las desventajas de la entrevista en profundidad.

12.5. Preparación y planificación de la entrevista

La entrevista no tiene reglas fijas y por lo tanto no se puede hacer siguiendo una receta; para muchos autores se trata de una práctica eminentemente artesanal, dependiente en gran medida —como acabamos de decir— del conocimiento personal y tácito del entrevistador. Combessie (1996/2005: 42) se refiere al "arte de la entrevista", y señala que un "análisis retrospectivo de [ella] revela que cada investigador tiene su propio 'estilo' de [conducirla]" derivado de "un saber práctico incorporado y de las propiedades sociales singulares de entrevistador y entrevistado". Alonso (1998: 77), por su parte, afirma que la entrevista es "refractaria a cualquier criterio cientificista de definición de la herramienta metodológica".

Sin embargo, esto no significa que se trate de una práctica totalmente anárquica. Por el contrario, hay una serie de cuestiones que pueden planificarse de antemano, aun reconociendo distintos grados de realización posible de las decisiones con anterioridad a la práctica misma. Esta planificación, o al menos reflexión sistemática previa acerca de los aspectos que a continuación trataremos, contribuye justamente a potenciar las capacidades artesanales del investigador en el marco de la entrevista.

La primera cuestión por resolver se refiere a la SELECCIÓN de los sujetos a entrevistar. Antes que nada debe quedar en claro que los criterios que generalmente se siguen difieren de aquellos aleatorios empleados en la determinación de una muestra a los fines del sondeo (véase capítulo 6). Esto al menos por dos motivos importantes:

a) las entrevistas en profundidad se emplean casi siempre en el marco de investigaciones que hemos llamado no estándar, que no tienen como objetivo principal la generalización estadística de sus resultados;
b) por las características mismas de la entrevista, la cantidad que se puede llevar a cabo en un estudio es reducida —habitualmente muy inferior al

necesario para que las ventajas de la extracción aleatoria resulten efectivamente plausibles.[4]

También es importante hacer notar que en lo concerniente a la selección de entrevistados no existe un criterio único, alternativo al de la aleatoriedad. Valles (1997: 213), siguiendo a Gorden (1969), propone tener en cuenta una serie de preguntas guía que pueden concebirse como consecutivas, a los efectos de restringir gradualmente el foco de los potenciales entrevistados:

1) ¿quiénes tienen la información relevante para la investigación?;
2) de ellos, ¿quiénes son más accesibles física y socialmente?;
3) ¿quiénes entre ellos están más dispuestos a cooperar brindando información al investigador?, y
4) entre los que cumplen todos los requisitos precedentes, ¿quiénes son los más capaces de comunicar la información de interés con precisión?

Para responder estas preguntas guía y tomar las decisiones de selección se recurre a una serie de estrategias diferentes que pueden emplearse individualmente o bien combinarse en el marco de una misma investigación. Tal vez la más conocida sea el MUESTREO INTENCIONAL: en este caso el investigador seleccionará a los entrevistados de acuerdo con un conjunto de criterios relevantes, que obviamente pueden cambiar de una investigación a otra. Frecuentemente se tienen en cuenta cuatro tipos de aspectos: la homogeneidad/heterogeneidad de la población de referencia desde el punto de vista de las cuestiones en estudio; la tipicidad/marginalidad/representatividad de los sujetos; las variables sociodemográficas (sexo, edad, nivel de estudios, nivel socioeconómico, etcétera) y otras cuestiones que permitan diferenciar a las personas sustantivamente en función del fenómeno de interés. A partir de la consideración de uno o varios de estos aspectos combinados, es frecuente segmentar a la población formando subgrupos de los cuales se seleccionará uno o más entrevistados con el fin de tener un panorama amplio de diferentes experiencias y discursos.

Cuando se trata de personas que por sus características no son fácilmente identificables por parte del investigador —pensemos por ejemplo en inmigrantes que conforman una minoría poco visible y la vez dispersa— se suele recurrir a la técnica de BOLA DE NIEVE (*snowball*): a partir de uno o unos pocos contactos iniciales, y valiéndose de las redes personales de los mismos, se busca ampliar progresivamente el grupo de potenciales entrevistados que compartan aquellas características que los hacen minoritarios o poco visibles y accesibles.

[4] Cabe recordar que el criterio de selección aleatoria no tiene sentido por sí solo, y debe darse junto con un cálculo de la cantidad de casos mínimamente razonable como para que sus ventajas se vuelvan operativas. Si pensamos en una población muy heterogénea, por ejemplo, el análisis de muy pocos casos seleccionados al azar de ningún modo garantizaría la calidad de los resultados en términos de una estimación de lo que efectivamente sucede en dicha población. En estas circunstancias, muy poca diferencia habría entre la selección de muy pocos casos al azar y de muy pocos casos por cualquier otro criterio.

Finalmente, las cuestiones de accesibilidad en general se relacionan con lo que se conoce como MUESTREO OPORTUNISTA (Burgess 1984): por medio de éste se seleccionan aquellos sujetos proclives a colaborar en el estudio y a los cuales el investigador tiene garantizado el acceso.

Una cuestión muy importante es aquella relativa al número de entrevistas que deben hacerse en el marco de una investigación. Para dar una respuesta a este problema se sigue habitualmente el criterio de la SATURACIÓN (Glaser y Strauss 1967):[5] se debería seguir realizando entrevistas hasta tanto se alcance la certeza práctica de que nuevos contactos no aportan elementos desconocidos con respecto al tema de investigación, o no emergen aspectos hasta entonces no tratados.

Otro preparativo clave de la entrevista consiste en el diseño de la GUÍA (a veces también llamada guión) que se utilizará en su desarrollo. Según Valles (1997: 219), se trata de "trazar un esquema, en el que se anticipen los modos de abordar el tema central y las cuestiones secundarias [...] supone tener listas preguntas de amplio espectro para los inicios, así como una serie de cuestiones y argumentos que sirvan (en caso necesario) para pasar de unos asuntos a otros; o para motivar al entrevistado". Sin embargo —como ya hemos indicado— su nivel de detalle puede variar, dando lugar a entrevistas más o menos estructuradas. En general existe una relación directa entre el grado de familiaridad con el tema y la capacidad de pensar de antemano la guía de entrevista: a mayor conocimiento, tanto del tema en cuestión como de los sujetos a entrevistar, mayores son las posibilidades de confeccionar una guía más completa y precisa. Sin embargo, aun en el caso de que se cuente con una guía detallada, debemos recordar que la entrevista no debe convertirse en un interrogatorio. La guía debe funcionar simplemente como un recordatorio, una ayuda instrumental que permita cubrir los temas relevantes de acuerdo con los objetivos de la investigación, pero sin imponer un orden determinado ni limitar de modo rígido las cuestiones a tratar. Como afirma Combessie (1996/2005: 37) "la entrevista debe seguir su propia dinámica", sin que la lista de temas contenida en la guía determine la formulación de las preguntas ni sus encadenamientos. Es común que ésta se construya a lo largo de un proceso por medio del cual el investigador va ganando familiaridad con el tema y con los entrevistados, a partir de lecturas específicas y de una sucesión de entrevistas exploratorias que permiten ir completando, modificando y refinando la guía, hasta alcanzar cierta convicción de que ella cubre todos los aspectos relevantes.

Aunque muchas veces tienden a desatenderse, los preparativos de la entrevista también deben incluir algunos aspectos relativos al investigador/entrevistador. Ya se ha indicado que sus características (por ejemplo género, edad, clase social, apariencia, etcétera) ejercen una influencia decisiva en los discursos que puedan emerger a lo largo de la conversación. Algunas de estas cuestio-

[5] Cabe señalar que Glaser y Strauss (1967) originariamente usan el término 'saturación' para referirse a las categorías que se construyen a partir de la información resultante de técnicas no estándar (por ejemplo, observación participante y entrevistas) en la fase de análisis (véase capítulo 17).

nes no son controlables/manipulables, y en el caso de que se consideren decisivas (en cuanto a un impacto negativo sobre los resultados), debería tenerse en cuenta la posibilidad de que otro miembro del equipo de investigación realice la/s entrevista/s. Pero hay ciertas cuestiones que sí se pueden pensar de antemano, especialmente la vestimenta, el tipo de lenguaje a utilizar, las formas de realizar el contacto y la presentación. Como se dijo en el apartado 12.3, la entrevista en profundidad se basa en un contrato de comunicación que implica saberes explícitos e implícitos compartidos. Entre estos últimos —se recordará— se encuentran los códigos lingüísticos y culturales, y las reglas sociales que hacen posible un determinado tipo de interacción; es responsabilidad fundamental del investigador adaptarse a los códigos del entrevistado, y esto es algo que se debe pensar —en la medida de lo posible— en la fase preparatoria de la entrevista.

Finalmente, hay otros preparativos que deben ser atendidos antes de realizar las entrevistas. Nos referimos especialmente a la selección del lugar en el que ésta se llevará a cabo y a la elección de la forma de registro de la conversación. El lugar no es un aspecto menor; un análisis retrospectivo de las entrevistas, e incluso de las conversaciones informales de la vida cotidiana, revela que los discursos pueden estar condicionados por el ámbito en el que se producen. Es muy posible que un adolescente —por ejemplo— diga cosas muy diferentes sobre un mismo tema si se encuentra en un local nocturno o en el patio del colegio junto a sus amigos, o en la casa familiar —bajo la presencia vigilante (físicamente evidente o no) de los padres. Por otra parte, el grado de formalidad, solemnidad, luminosidad, etcétera, de un ambiente incide en la situación de entrevista, y muy probablemente no sólo en términos de la comodidad física del entrevistador y del entrevistado. Abundan los ejemplos de entrevistados que tuvieron dificultad para articular un discurso continuo en un espacio físico poco familiar y percibido como hostil; también hay registro de situaciones en las que el espacio ayudó al entrevistado a revivir cierto tipo de experiencias clave para el relato; pero se debe tener en cuenta que si éstas hubieran sido traumáticas, el efecto de entrevistar al sujeto en ese ambiente podría ser estresante y por lo tanto no adecuado.

En cuanto al registro de la entrevista, en la medida en que el entrevistado dé su consentimiento éste se suele hacer recurriendo al grabador. De este modo no se pierden detalles de la interacción verbal, y el entrevistador puede concentrarse en el desarrollo de la conversación, siguiendo atentamente su hilo y registrando también el lenguaje gestual. Como indica Ferrarotti (1986/1991: 19): "no hablan sólo las palabras, sino los gestos, las expresiones del rostro, los movimientos de las manos, la luz de los ojos". Si el investigador se concentra en el registro manual del relato, entonces pierde la oportunidad de sensibilizarse frente a estos otros aspectos de la comunicación.

Muchas veces se plantea que la grabación genera cierto tipo de rechazo por parte de los entrevistados y, en el mejor de los casos, una considerable reactividad. Pero no es seguro que la presencia de una persona —en este caso el entrevistador— que escribe todo lo que se dice en la entrevista no sea igualmente reactiva, con la desventaja —como apuntábamos más arriba— de que se pierden otros detalles de la interacción y se interrumpe la fluidez de la conversación. Por otra parte, la tecnología actual de grabación digital permite que los dispositivos se vuelvan prácticamente inconspicuos, al punto de que habitualmente pocos minutos después del inicio de la entrevista el entrevistado se olvida de su presencia.

A pesar de todos los preparativos que puedan hacerse, es importante recordar —como ya se ha dicho en varias oportunidades— que en la situación de entrevista la relación entre entrevistador y entrevistado es en cierto sentido imprevisible. Ferrarotti (1986/1991: 19) afirma que se trata de "una relación verdaderamente humana, es decir dramática, sin resultados asegurados". La clave de su éxito radica en las formas de interacción verbal, y éstas descansan sobre saberes tácitos que recurrentemente exigen la puesta en juego de lo que Valles (1997) llama "tácticas del entrevistador": conocimientos que forman parte de su oficio y pericia y que improvisa durante la realización de la entrevista según el curso que vaya adquiriendo la interacción verbal. Para Alonso (1998) tal interacción se establece inicialmente a partir de la intervención del investigador, por medio de consignas y comentarios, y de ahí en más a través de actos del habla que pueden clasificarse como declaraciones, interrogaciones y reiteraciones. Combessie (1996/2005: 39) destaca el recurso a los relanzamientos: "la repetición por parte del entrevistador de una parte (una palabra, un elemento de la frase, una frase) de lo que se acaba de decir". Esto pone en evidencia la atención del entrevistador por el relato del entrevistado y "provoca precisiones suplementarias, confirmaciones o reformulaciones de su discurso". Valles (1997: 220-221), por su parte, destaca entre las tácticas del entrevistador el silencio, la animación y elaboración, la reafirmación y repetición, la recapitulación, la aclaración y el cambio de tema. Finalmente, concede un lugar especial a la táctica de la post-entrevista: una prolongación de la conversación luego de que ésta haya terminado desde el punto de vista formal. En ella se suele producir "una cierta redefinición de la situación y de los *roles* respectivos", dando lugar a la emergencia de cierto tipo de discurso —a "micrófono cerrado"— que deriva muchas veces en la recolección de informaciones que el entrevistado no ha explicitado durante la entrevista formal.

CAPÍTULO 13

FOCUS GROUP Y OTRAS FORMAS DE ENTREVISTA GRUPAL

13.1. El *focus group*

Las entrevistas grupales se caracterizan por la presencia simultánea de varios entrevistados que se relacionan a través de técnicas conversacionales. El FOCUS GROUP es un tipo de entrevista grupal que se enmarca dentro de los enfoques cualitativos; su dinámica se basa en organizar un grupo particular de personas para que discutan un tema determinado que constituye el objeto de la investigación. La información, producto de esa interacción que refleja las perspectivas y experiencia grupales, es registrada y clasificada para su análisis. Finalmente, el material se analiza para elaborar conclusiones y sugerencias.

Los miembros del grupo son seleccionados en función de los objetivos del estudio y de acuerdo con ciertos requisitos que la técnica establece; generalmente se trata de individuos que comparten ciertas características (sexo, edad, nivel económico-social). Los grupos se conforman de manera homogénea según estos criterios y son heterogéneos entre sí, dentro del conjunto de grupos estudiados. Esta heterogeneidad permite realizar un análisis comparado entre los diferentes grupos que han sido diseñados en el marco de una investigación.

La técnica ha tomado diferentes denominaciones: grupos focales, grupos focalizados, grupos de discusión y grupos motivacionales, denominación que alude a la búsqueda de factores no observables que surgen en el transcurso de la conversación. En realidad, se trata de una entrevista múltiple donde se promueve la interacción entre los entrevistados, en la medida que la información buscada no son las actitudes y opiniones individuales sino el resultado de la situación social de debate, donde unas opiniones son influidas por otras. En este sentido, las preguntas son presentadas en forma abierta al grupo en su conjunto y se estimula a los participantes a intercambiar ideas con los otros miembros.

Robert Merton (1946) fue uno de los primeros en establecer los lineamientos de los *focus groups* en su artículo "La entrevista focalizada". La técnica se desarrolló durante los años cincuenta en la evaluación de audiencias de radio (Stewart y Shamdasani, 1990); posteriormente, la introducción de tecnología

sumó a los grupos presenciales conferencias telefónicas (a fines de los sesenta) y grupos *on line* en los años noventa, lo que permitió la integración de grupos con individuos distantes.

Un estudio basado en grupos focalizados requiere: una selección adecuada de los integrantes de los grupos, un moderador/facilitador entrenado, un clima de confianza y tolerancia, y analistas experimentados. Con estos elementos se pone en marcha un proceso que provee una gran cantidad de valiosa información en corto tiempo.

El *focus group* se caracteriza por ser muy pautado en su diseño y estructura a fin de garantizar la calidad de la información y el control del tiempo, y por la baja estructuración en los contenidos de la información que fluyen libremente pues sólo son orientados a fin de evitar que los participantes se desvíen de la temática propuesta. Es decir que, dentro de un proceso minuciosamente diseñado, los datos se producen en forma espontánea en la dinámica de la interacción grupal. La relevancia de esta técnica está centrada en la comunicación entre los entrevistados, ya que es en el proceso comunicacional donde se construyen las matrices de significado social.

13.2. Objetivos

El grupo focalizado se desarrolla en un escenario que permite observar todos los elementos que se ponen en juego en la discusión grupal: actitudes, emociones, creencias, opiniones, experiencias, evaluaciones, reacciones, consensos y disensos. Estos elementos serán articulados luego —a través del análisis— en estructuras de sentido.

A diferencia de la entrevista individual, dirigida al estudio de actitudes y opiniones de cada persona entrevistada, el grupo focal se orienta hacia una pluralidad de perspectivas vinculadas a una temática, que surge en la interacción colectiva, en cuya dinámica se ponen al descubierto las motivaciones y cosmovisiones de los participantes así como los usos frecuentes del lenguaje. En el intercambio de ideas los sujetos pueden reconsiderar sus propias posiciones y, como producto de la comunicación, suelen surgir temáticas y perspectivas que no estaban explícitas con anterioridad a la discusión. Este tipo de información no es recabable a través de otras técnicas de recolección de datos como la observación, la entrevista en profundidad o la encuesta.

Si bien el grupo focalizado se desarrolla en un escenario organizado y preestablecido donde todo (incluido el propio grupo) ha sido construido a tal efecto, la utilización de preguntas abiertas y el análisis interpretativo de los datos lo ubican dentro de las perspectivas no estándar (véase apartado 2.5). De modo que, metodológicamente, no está orientado hacia la medición sino hacia la comprensión de conductas y actitudes.

La principal riqueza de la técnica, entonces, reside en la reformulación de los significados que se produce en la dinámica grupal. En este sentido, los grupos focalizados pertenecen a un tipo de investigación interpretativa donde el sujeto de interpretación es un grupo social intencionalmente construido.

13.3. Utilización

En sus inicios, los grupos focalizados se utilizaron en estudios de mercado, donde continúan teniendo mucha aplicación para el estudio de actitudes en relación con productos y servicios, evaluación de conceptos, publicidad y envases (Krueger 1991). En forma más reciente se aplicaron en estudios académicos, en particular en la medicina (Powell & Single 1996) y la sociología. Su utilización en la política se ha orientado a los estudios de opinión pública, preelectorales y de evaluación de gestión o de políticas públicas, y en las investigaciones de comunicación se aplican en los estudios de audiencias.

Resultan particularmente útiles en estudios de evaluación y como campo de recolección de fragmentos de discurso o conceptos a ser utilizados en campañas publicitarias o electorales. Se realizan en diferentes etapas de las campañas: al inicio en la búsqueda de insumos discursivos y para la toma de decisiones estratégicas; en etapas intermedias indagando efectos y evaluaciones, y al final del proceso de la campaña, para evaluar el impacto de ésta.

Se pueden utilizar como un método en sí mismo o en combinación con otros métodos de investigación. Cuando se incluyen en estudios multimétodo, se utilizan tanto dentro de estudios exploratorios previos a una medición, a fin de indagar tendencias a ser corroboradas o descubrir nuevas dimensiones del análisis, como en etapas posteriores dirigidas a profundizar la comprensión de ciertas tendencias observadas. También pueden combinarse con todo tipo de técnicas no estándar (también llamadas "cualitativas"), como entrevistas individuales en profundidad (véase capítulo 12), análisis documental, biográfico, etcétera. Igual que en las entrevistas individuales, sus observaciones y conclusiones no pueden generalizarse a poblaciones más amplias, ya que los sujetos que los integran no constituyen una muestra estadísticamente representativa.

13.4. La dinámica de la sesión grupal

Los elementos necesarios para poner en marcha un grupo focalizado son: un espacio físico adecuado, la presencia de los integrantes del grupo, al menos un moderador grupal y equipos para el registro de la información. Si no se cuenta con equipos de grabación, se recomienda la presencia de dos moderadores, uno de los cuales tendrá como única función tomar notas de campo.

La selección de la locación se deberá realizar teniendo en cuenta los objetivos del estudio y las características del grupo seleccionado; de acuerdo con estos criterios variará el grado de formalidad del lugar, así como su cercanía. Pero, en todos los casos, se debe tratar de un lugar confortable que dé cabida a diez personas, que todas puedan verse simultáneamente (se recomienda situarlos alrededor de una mesa redonda), que invite a la conversación y no tenga elementos que podrían resultar inhibitorios u orientar las opiniones. En este sentido se recomienda que la locación constituya un lugar neutral (Krueger 1991).

Cada grupo se integra en forma homogénea según ciertas características de sus integrantes que los convierte en miembros de algún segmento poblacional,

de tal modo que sus verbalizaciones puedan ser interpretadas como opiniones, actitudes, ideas y evaluaciones presentes en ese segmento social. Este requisito cumple dos funciones: una vinculada al objetivo de la técnica, en la medida en que los resultados de la discusión se considerarán como expresiones correspondientes al segmento de población seleccionado, y la otra relacionada con su dinámica, ya que la similitud social entre los integrantes facilitará el diálogo y la participación. La información recabada no es considerada como dato individual, tampoco como dato generalizable o cuantificable, sino como expresión grupal de opiniones existentes en determinados segmentos sociales.

El tamaño de los grupos varía entre 6 y 10 personas. Éstas pueden ser invitadas a participar en forma personal, por teléfono, por correo postal o por correo electrónico. Se recomienda confirmar la participación con anterioridad a la reunión para garantizar la presencia de los entrevistados. La selección de los participantes es intencional y se puede realizar a través de reclutadores profesionales, de informantes clave o de redes sociales u organizacionales.

Un requisito del diseño es que la gente que conforma un grupo no se conozca con anterioridad, a fin de que la dinámica grupal no se contamine con relaciones previas y porque la diversidad de individuos se traduce en pluralidad de posiciones respecto del tema en discusión (este requisito puede no cumplirse si lo exige el diseño del estudio).[1]

Una sesión grupal dura entre hora y media y dos horas, durante las cuales el moderador deberá mantener la discusión focalizada en el tópico en cuestión. Cuando los temas se dispersen deberá reencauzar el debate con preguntas estratégicas estimulando la participación de todos los integrantes en forma equilibrada. La intensidad de la interacción produce, en este corto tiempo, un flujo importante de información.

La estimulación a la discusión se realiza a través de preguntas abiertas, y es muy importante la participación de todos los sujetos pues todas las voces deben ser expresadas durante la sesión. El moderador debe evitar el predominio de algunas voces que puedan intimidar a otros participantes, a través de la invitación a hablar a los más silenciosos.

Algunos salones utilizados en la investigación de mercado están equipados con *cámara Gesell*, un espacio especialmente diseñado para realizar observaciones a través de una ventana espejada o espejo unidireccional que permite la observación en un solo sentido. Desde allí los investigadores y los clientes pueden observar el trabajo del grupo sin ser vistos por sus integrantes.

A diferencia de la entrevista individual, en los *focus groups* no es posible garantizar el anonimato ni la confidencialidad de los datos, ya que las opiniones son vertidas en presencia de los miembros del grupo. Asimismo, a diferencia de las técnicas no estándar en general, el escenario del grupo focal se caracteriza por no ser natural sino construido y organizado por el equipo de investigación. Se reúne a un grupo de gente desconocida, en un lugar que le es

[1] Por ejemplo, si se aplica esta técnica dentro de una empresa y los grupos se conforman con miembros de diferentes departamentos de ésta.

ajeno y se le solicita que intercambien ideas sobre un tema que les es propuesto. Una de las críticas posibles a esta técnica señala la posibilidad de que la artificialidad de esta situación, resultado de la descontextualización de los entrevistados, distorsione los discursos y las discusiones.

Cuando se contrata a reclutadores profesionales se pueden producir problemas que afectan la validez de los grupos, como la presencia de los denominados "respondentes profesionales", quienes participan de grupos focalizados como una actividad regular.

La dinámica de la sesión de un grupo focalizado se puede dividir en tres momentos. El primero consiste en la presentación del grupo, donde el moderador explica los objetivos de la reunión, las reglas de la interacción y el uso del tiempo, además de presentar a los participantes. El segundo momento consiste en el planteo de las preguntas abiertas que darán lugar a las intervenciones y la promoción del debate. Es recomendable comenzar por preguntas muy generales que permitan instalar la temática para luego ir precisando el foco con preguntas cada vez más específicas. El último momento corresponde al cierre, donde el moderador agradece la participación, puede sintetizar las etapas posteriores de la investigación y se despide. Habitualmente se ofrece un incentivo o presente para estimular la aceptación a participar.

13.5. Tipos de *focus group*

Existen tres tipos de grupos focalizados: los tradicionales o cara a cara, los grupos en teleconferencia o telefónicos y los grupos *on line*, a través de Internet.

A) Grupos cara a cara

Los grupos tradicionales son los más utilizados y su principal ventaja es la presencia simultánea de los entrevistados en un mismo espacio. Asimismo, cualquier persona puede ser invitada a participar, a diferencia de los otros tipos de grupos focales. La presencia simultánea en un mismo espacio físico de los integrantes del grupo permite que se genere una gran cantidad de datos que trascienden el discurso verbal, como la gestualidad y el lenguaje corporal que expresan emociones difíciles de captar cuando la presencia es remota. Sin embargo, estos elementos visuales sólo se registran en el caso de grabar al grupo en *videotape* o cuando existe una moderación doble donde uno de los coordinadores toma nota de estas expresiones. Otra de sus ventajas es la posibilidad de presentar todo tipo de materiales que pueden ser evaluados tanto visual como auditivamente y por sus texturas. Sus principales desventajas residen en tener que reclutar sujetos dentro de un área geográfica limitada y que éstos deben desplazarse hasta el lugar de reunión, ya que esto incide en los costos y también en la tasa de aceptación a participar. Por otro lado, la desgrabación y transcripción del material necesarias para su análisis demoran algunos días.

B) Grupos telefónicos

Los grupos focalizados en teleconferencia se desarrollaron hacia fines de la década de 1960. Su utilidad frente a los grupos cara a cara estriba en la posibilidad de integrarlos con individuos que residen en diferentes lugares geográficos. Asimismo, resultan exitosos para reunir personas que por su estatus o características no asistirían a un grupo cara a cara, como altos ejecutivos de empresas, científicos, políticos, artistas, etcétera. Entre sus ventajas se señalan también el menor costo, que los integrantes no necesitan trasladarse de su ambiente natural, un mayor grado de anonimato, ya que los participantes no se conocen, y que las transcripciones son accesibles en tiempos muy cortos. Entre las principales desventajas se cuentan la imposibilidad de presentar materiales, aunque éstos pueden ser enviados con anterioridad por correo, fax o a través de Internet, y que sólo pueden participar personas con acceso al teléfono.

Para realizar grupos focales telefónicos es necesario contar con un sistema de teleconferencias telefónico. Es importante tener en cuenta que cuanto mayor es el uso de tecnología, más restringida es la población accesible a través de estos métodos.

C) Grupos focales *on line*

La incorporación de Internet a la técnica de los grupos focales es muy reciente: data de la década del noventa. Para su implementación es necesario que los participantes tengan acceso a una computadora y a Internet. Esto reduce en gran medida el universo accesible, que se limita a individuos de mayor nivel económico y educativo. Una desventaja adicional en este sentido es la baja tasa de aceptación a participar.

Al igual que los grupos telefónicos, tienen la ventaja de acceder a individuos dispersos sin necesidad de traslado, lo que incide positivamente en la organización logística y los costos. Por otro lado, presentan la ventaja de poder mostrar todo tipo de materiales e imágenes a través de Internet.

A diferencia de los grupos cara a cara, no se produce una dinámica interactiva fluida, ya que frente a cada pregunta todos los sujetos deben tipear simultáneamente su respuesta. De modo que éstas se producen en forma paralela y esto resiente el flujo de influencia mutua entre los participantes. Una de sus principales ventajas es que la transcripción del material se produce durante la misma sesión.

13.6. La coordinación

La actividad del coordinador es clave para el éxito de esta técnica. Su rol consiste fundamentalmente en promover el interés, la participación, la interacción y el debate entre los entrevistados. Su objetivo último es lograr que el grupo funcione con autonomía, permitir que las ideas fluyan libremente y sólo intervenir para guiar la discusión manteniendo focalizado su contenido y estimular a todos los miembros a participar.

El coordinador trabaja con una pauta de entrevista organizada en base a una lógica que va de lo general a lo particular. Las preguntas abiertas iniciales son más generales y permiten ubicar al grupo en la temática a tratar; enseguida se acerca la discusión al objetivo deseado a través de preguntas más puntuales. Es tarea del moderador mantener el foco de la discusión sin inhibir la participación ni orientar las respuestas y guiar la interacción de modo que las voces dominantes en el grupo dejen espacio a los que permanecen en silencio. Todos los participantes deben tener la oportunidad de expresarse y el moderador debe promover que así ocurra.

Es importante destacar que el moderador no forma parte de la discusión, no plantea posiciones ni perspectivas y tampoco juzga las intervenciones; su tarea consiste en generar un clima de confianza para la interacción y en mantenerla enfocada en el tópico de la investigación. La mayor dificultad de la coordinación es estimular a los participantes para que profundicen en el foco de discusión, de modo que surjan las motivaciones de las opiniones y los factores subyacentes que funcionan como otorgadores de sentido. Con este fin, el moderador puede utilizar diferentes técnicas para estimular las asociaciones, como juegos de presentación, tests proyectivos y *role playing*.

Para una realización eficiente del grupo, es recomendable que el moderador sea una persona entrenada en la dinámica de grupos y tenga ciertas habilidades personales, como saber escuchar, no prejuzgar, manejar los silencios y saber mantener, en todo momento, una actitud neutral generando un clima de confianza que estimule la participación.

13.7. El registro y análisis de la información

En los grupos cara a cara, habitualmente, la información se registra a través de una grabación de audio. Sin embargo, resulta interesante grabarla en *videotape*, ya que la imagen permite observar las expresiones de lenguaje no verbal. Posteriormente, la información es desgrabada y transcripta para su análisis. La transcripción debe ser textual y sin omisiones, dejando incluso registro de los silencios. En los grupos con doble coordinación, uno de los moderadores toma notas de campo intentando capturar elementos o características que aporten al clima de la discusión y que no quedarán registrados en el grabador; estas notas forman parte del material de análisis.

Después de desgrabada la información, ésta es procesada a través de su codificación y clasificación. Una vez así organizada, se procede a su análisis en dos niveles: la narración descriptiva y la interpretación. La codificación de los textos, que supone principios de clasificación basados en los criterios de relevancia de la investigación, permite su posterior análisis. Ésta puede hacerse sobre la propia trascripción o a través del uso de grillas y matrices; sin embargo, cada vez más se utilizan programas informáticos para el tratamiento de los textos (ATLAS ti, NUD*IST, etcétera). La información codificada y clasificada se organiza en esquemas de categorías conceptuales que le otorgan sentido.

A través de la descripción se reproducen en forma organizada los discursos

de los integrantes; se describe lo expresado en el grupo sin identificar a los individuos, señalando las opiniones y creencias predominantes o consensuadas grupalmente, así como los disensos. Se debe evitar el uso de cantidades o porcentajes, ya que no existe una medición en sentido estricto, y utilizar evaluaciones cualitativas escalares como: algunos, todos, la mayoría, pocos. Finalmente, a través de la interpretación, se reorganiza la información generando nuevos núcleos de sentido y se preservan ciertos fragmentos del discurso que serán incorporados en forma textual y encomillados al informe final, para ilustrar los resultados del proceso interpretativo.

Como en todo análisis cualitativo, la habilidad y subjetividad del analista influyen en forma peculiar en los resultados. Es muy importante, entonces, quién analiza los datos. Lo ideal es que el moderador, quien realiza las transcripciones y el analista sean la misma persona. De este modo son mayores las probabilidades de que se preserve el clima predominante en la discusión del grupo y se pierda menos información. Pero no siempre el trabajo se realiza de este modo: a menudo estos roles son desempeñados por diferentes personas especializadas, en particular en la investigación de mercado. En este caso, es importante que el responsable del análisis tenga la posibilidad de tener contacto con los moderadores grupales.

Todos los pasos del *focus group* deben reflejarse en el informe final. Éste debe contener los objetivos del estudio y la estrategia metodológica, así como toda la información relevante de la puesta en marcha de los grupos focales en sus aspectos formales: número de grupos, número de participantes, criterios y procesos de organización y selección, fechas de realización, cantidad de moderadores, tipos de registro, etcétera. En el informe no es posible incluir toda la información registrada de los contenidos, pero es importante que aparezcan todas las opiniones y posiciones expresadas.

13.8. Otras formas de entrevista grupal

Existen otros tipos de entrevistas grupales; en algunas el énfasis no está puesto en la interacción grupal —como en el grupo focal—, sino en las respuestas individuales de cada uno de los integrantes. En estos casos es menos importante que los miembros del grupo no se conozcan con anterioridad, y es habitual que se apliquen a grupos que tenían una existencia previa dentro de una institución o comunidad (sindicato, partido político, organización social). Las recomendaciones y la dinámica de estas entrevistas son semejantes a las de la entrevista individual (véase capítulo 12).

También es posible realizar una entrevista a miembros de un grupo en forma simultánea con un cuestionario autoadministrado. En este caso se utiliza un instrumento estructurado que cada miembro del grupo responde por escrito en forma individual; no existe interacción entre los entrevistados y sus respuestas son procesadas con las técnicas de la encuesta.

Otros tipos de entrevistas grupales se diseñan con objetivos diferentes a los expuestos. La tormenta de ideas o *BRAINSTORMING*, originada en la década de

1950 en el área de la publicidad, a partir del libro de Alex Osborn *Applied Imagination*, se orienta fundamentalmente a la generación de nuevas ideas y se distingue por su carácter inestructurado y exploratorio. Se la utiliza particularmente en la búsqueda de solución de problemas, a través de la interacción creativa de un grupo de expertos o especialistas o de un grupo de trabajo. La idea central es generar nuevas perspectivas para enfocar un problema que den lugar a soluciones innovadoras que permitan tomar mejores decisiones. Se recomienda a los participantes que expresen sus ideas independientemente de su viabilidad o razonabilidad, sin inhibiciones, sin importar cuán descabelladas resulten o parezcan, ya que en el intercambio grupal se intentará convertirlas en caminos posibles y útiles. Dado que cuanto mayor sea el número de ideas expresadas mayores serán las probabilidades de llegar a una solución, toda idea es aceptada en forma acrítica, focalizando en sus potencialidades. El supuesto que subyace a esta técnica es la posibilidad de enriquecer grupalmente las ideas individuales. Tiene en común con la dinámica del *focus group* la búsqueda de un producto grupal y la presencia de un moderador, quien mantiene focalizada la temática de la discusión y promueve la participación de todos los miembros del grupo.

El GRUPO TERAPÉUTICO constituye otra forma particular de entrevista grupal orientada a la elaboración colectiva de problemas que afectan la vida individual. Se relaciona con la búsqueda de ayuda a través de procesos emocionales que se desatan o expresan en la interacción grupal. El grupo es liderado por un terapeuta especializado, en quien reside la responsabilidad de la conformación del grupo, la definición última de las reglas que guiarán las sesiones y la interpretación de los emergentes grupales. Los contenidos son abiertos basados en la libre asociación. Tiene diferentes dinámicas en función de la perspectiva teórica que lo guíe y se caracteriza porque las reuniones se repiten y se prolongan en el tiempo con un horario estable.

Los denominados grupos nominales o TÉCNICA DELPHI se originan también a comienzos de la década del cincuenta, para el tratamiento de problemas estratégicos basado en las opiniones de expertos. La idea central es consultar a un grupo de especialistas sobre determinado tema y circular las respuestas en reiteradas rondas para lograr un consenso. Entre los entrevistados no existe interacción simultánea o presencial, pero cada uno toma conocimiento de las respuestas de los demás en forma anónima y, a partir de éstas, puede reformular sus propias respuestas, de tal modo que las diferentes opiniones se van modificando en el proceso iterativo y tienden hacia la convergencia. Los cuestionarios utilizados están conformados por preguntas muy precisas que dan lugar a respuestas cuantificables, a las cuales se aplican estadísticos como medidas de tendencia central y dispersión. A través de los cuestionarios sucesivos se tiende a precisar estas medidas. Habitualmente el Delphi consta de cuatro circulaciones. En las sucesivas rondas se dan a conocer los estadísticos calculados y se indaga a los entrevistados sobre su acuerdo con la respuesta media obtenida, y así se va delineando la posición del grupo. Se recomienda no utilizar más de 30 expertos. La técnica se puede aplicar a través de correo electrónico y también existen programas que permiten la aplicación de *Delphi on line*. La técnica

Delphi comparte con el *brainstorming* una idea de agregación, que toma en cuenta la participación de todos los entrevistados para dar lugar a un resultado nuevo y colectivo, pero se distingue en que se trata de un método pautado, procesado con técnicas cuantitativas que conducen a una salida consensuada o a destacar las divergencias.

CAPÍTULO 14

ESTUDIO DE CASO/S

14.1. ¿Método o estrategia de investigación?

El ESTUDIO DE CASO (EC) se basa en un diseño de investigación orientado al análisis de las relaciones entre muchas propiedades concentradas en una sola unidad. Trabajar con muchas variables y un caso (o muy pocos) ha dado lugar a diversas estrategias de abordaje encaminadas a la búsqueda de múltiples fuentes de evidencia (Yin 1993).

Algunos autores definen el EC a partir de N=1 (Lijphart 1971, Eckstein 1975, Bartolini 1991), pero otros incluyen dentro de estos estudios la posibilidad de trabajar con más de un caso (Stake 1994, Yin 1993)

Stefano Bartolini (1991) lo grafica como una fila de la matriz de datos cuando se trata de un estudio sincrónico (un caso observado en un corte en el tiempo) y lo diferencia del 'estudio de desarrollo de caso' cuando incorpora el tiempo como una dimensión de variación. El autor hace esta distinción a partir de la obra de Eckstein (1975), quien no admite el estudio de caso en una variación temporal, caracterizando a los estudios diacrónicos de casos como análisis comparativo. Sin embargo, Bartolini sostiene que un caso estudiado en una dimensión temporal de variación normalmente no se analiza de acuerdo con las reglas del método comparativo.

Por otro lado, Giovanni Sartori (1990) considera el análisis comparativo y el estudio de caso como estrategias complementarias, y sostiene que, si bien los estudios de caso son implícitamente comparativos, cuando se trata de un solo caso no se aplica el método comparativo.

Otros autores como Stake (1994) y Yin (1993) incluyen en sus categorizaciones del EC la posibilidad de analizar más de un caso en forma comparada. El primero los denomina 'estudio de caso colectivo' y el segundo se refiere a 'casos múltiples'. Sin embargo, si bien aceptan ampliar el diseño del EC haciéndolo extensivo a una multiplicidad de unidades, no hacen una referencia explícita al desarrollo del caso en el tiempo. Sus análisis parecen aludir más a estudios comparativos sincrónicos o replicaciones que al "estudio de desarrollo de caso" propuesto por Bartolini.

En algunos trabajos al EC se lo considera un método (Lijphart 1975; Yin

1993; Flyvbjerg 2004) a pesar de que resulta difícil señalar características del EC que permitan diferenciarlo de otros métodos de investigación, asignándole la autonomía necesaria para hablar de un método diferente. Parece más adecuado considerarlo una estrategia o un diseño de investigación, al cual generalmente se hace referencia en vinculación con los métodos "cualitativos" (Feagin, Orum y Sjoberg 1991; Stake 1994) o con el método comparativo (Lijphart 1971; Eckstein 1975; Sartori 1990).

Consideramos adecuada la propuesta de Stake (1994: 236) cuando sostiene que el EC no se trata de una opción metodológica sino de la elección de un objeto de estudio; es el interés en el objeto lo que lo define y no el método que se utiliza. Cualquier unidad de análisis puede convertirse en ese objeto ("el caso"), el cual se puede tratar tanto de una unidad individual como colectiva —una persona, una institución, una empresa, un programa, una política, una comunidad, un sistema, un país, etcétera. Una vez definido el objeto, en él se concentra toda la atención investigativa orientada a un análisis intenso de sus significados con la intención de comprenderlo en su especificidad más que buscando generalizaciones. El objeto se puede abordar desde diferentes métodos y con diversas técnicas de recolección de datos y análisis (observación, entrevistas en profundidad, análisis documental, análisis de datos estadísticos, sondeos, etcétera).

El predominio de las perspectivas no estándar en los EC se basa en diversos factores:

a) tratándose de una o pocas unidades, el enfoque predominante en los EC es el análisis en profundidad; la búsqueda no se orienta hacia el establecimiento de regularidades empíricas sino hacia la comprensión del caso en su unicidad;
b) resulta difícil realizar inferencias generalizantes a partir de una sola (o muy pocas) unidades, y
c) los estudios se llevan a cabo, en la mayoría de los casos, en el escenario natural donde los actores se relacionan dando lugar a los fenómenos que se intenta estudiar.

Sin embargo, a pesar del predominio de los métodos no estándar (o "cualitativos") y de los enfoques naturalistas, holistas o fenomenológicos en la mayoría de las investigaciones basadas en EC, en su diseño no se excluye la aplicación de métodos estándar (Stake 1994), cuando se intenta abordar el caso en sus dimensiones más estructurales o macro. Yin (1993) destaca explícitamente que el EC no implica ninguna forma particular de recolección de datos, ésta puede ser tanto cuantitativa como cualitativa.

En resumen, metodológicamente los EC son, generalmente, de tipo multimétodo. Dado que el objetivo es abordar un fenómeno complejo en forma holística, el investigador se aproxima al caso a través de diferentes métodos de investigación o triangulación metodológica (véase el apartado 2.4). El análisis casuístico no aparece, entonces, confinado a los límites de ningún método en particular; constituye un buen ejemplo del despliegue de estrategias a las que

Denzin y Lincoln se refieren como *bricolage*.

> "El *bricoleur* produce un *bricolage*, esto es, un conjunto de piezas unidas, un tejido de prácticas que dan soluciones a un problema en una situación concreta. 'La solución (*bricolage*), como resultado del método del *bricoleur*, es una construcción [emergente]' (Weinstein & Weinstein 1991: 161) que cambia y toma nuevas formas a medida que se le agregan nuevas herramientas, métodos y técnicas al rompecabezas" (Denzin y Lincoln 1994: 2-3).

Los antecedentes de EC más sistemáticos pueden encontrarse en los trabajos de los autores de la Escuela de Chicago y en la propuesta metodológica de Znaniecki (1934) sobre la INDUCCIÓN ANALÍTICA, que permitiría, a partir de un caso, iluminar aspectos de la teoría general, en oposición a la inducción enumerativa.

El posterior predominio de la Universidad de Columbia y de la investigación basada en sondeos dio lugar a fuertes críticas epistemológicas a los EC, particularmente centradas en sus debilidades frente a los requisitos estándar de rigor científico (véanse los apartados 1.1, 2.2, 3.2). Las principales críticas se focalizaron en dudas acerca de su carácter científico, dada la incapacidad para generalizar y construir teoría a partir del estudio de un caso único. Su validez y confiabilidad fueron cuestionadas, y el diseño en su totalidad quedó subordinado a las etapas exploratorias de investigaciones estándar más amplias.

Un nuevo interés por los EC comenzó a desarrollarse durante los años sesenta, en el marco de la revalorización de los enfoques cualitativos. Al calor de los debates epistemológicos de ese momento, orientados a modificar o relativizar las rígidas definiciones positivistas de la cientificidad imperantes, nacieron nuevas y renacieron viejas perspectivas tendientes a reconocer el carácter científico y las posibilidades generalizadoras de los EC. Su importancia estriba en que el diálogo entre el investigador y el caso inmerso en la vida real lo convierte, en alguna medida, en el diseño paradigmático dentro de las perspectivas que priorizan el estudio de los fenómenos sociales contextualizados.

La revalorización científica del EC aparece expresada en la obra de Campbell, quien, en la década de 1960, consideró que carecía de todo valor científico (Campbell y Stanley 1963), y posteriormente en los setenta modificó su posición y aceptó que los EC tienen su propio rigor y que ofrecen base empírica para falsear generalizaciones (Campbell 1975). Algunos autores consideran que, si bien los EC no permiten producir generalizaciones, un caso negativo puede establecer límites o modificaciones a ellas (Stake 1994) o debilitar una teoría (Lijphart 1971), mientras que otros atribuyen a la teoría el papel de generalizar los resultados a otros casos similares (Yin 1993).

Sin embargo, otras propuestas consideran que a través de esta estrategia se genera una base empírica para conceptos y generalizaciones, convirtiéndola en un elemento esencial para la investigación social (Feagin 1991), de tal modo que no sería correcto afirmar categóricamente la imposibilidad de generalizar a partir de un solo caso, en la medida en que esto es posible cuando se considera co-

mo CASO EJEMPLIFICADOR. Dentro de esta orientación, el concepto y las teorizaciones sobre el panóptico foucaultiano constituyen un ejemplo de generalización a partir de estudios de casos (Flyvberg 2006).

Desde una perspectiva cualitativa, el valor científico del EC estriba en su carácter de estudio denso, narrado en toda su diversidad a fin de desentrañar sentidos generales, metáforas, alusiones, alegorías que se expresan a través de múltiples marcas en la unicidad del caso. Si bien no provee elementos de prueba de enunciados generales, provee muchos indicios que pueden considerarse de apoyo a éstos. Además, puede constituirse en fuentes conceptuales, categoriales y de bases de datos inherentes a la construcción de teoría, así como en un elemento crítico importante en el proceso de reformulación de generalizaciones, al estilo del experimento crucial de Bacon y del falsacionismo de Popper; en particular ante la coexistencia de teorías rivales.

14.2. Cuándo hacer un estudio de caso

Generalmente se diseña un estudio de caso ante un objeto particular que despierta el interés del investigador o por la decisión de elegir un objeto de estudio típico para dar cuenta de un fenómeno a observar.

En algunas oportunidades el objeto de estudio no es el resultado de una elección sino que llega al investigador ya caracterizado como tal; un ejemplo es el caso terapéutico (se ha sostenido que el término 'estudio de caso' tiene su origen en la medicina). La mayoría de las veces, el médico aborda la investigación de un paciente en su particularidad y, también, dentro del marco de protocolos cuantitativos de tipo estadístico construidos a través de mediciones de síntomas y respuestas a estímulos. El paciente es estudiado en forma intensiva en su individualidad, contextualizado en sus condiciones presentes con referencia a su historia clínica y familiar, pero también comprendido dentro de márgenes provistos por los estudios estadísticos. Stake (1994) llama estudio de caso intrínseco a este tipo de investigaciones que son motivadas exclusivamente por el interés en las peculiaridades del objeto.

Las siguientes son algunas de las situaciones que dan lugar al diseño de un EC:

1) La aparición de un fenómeno nuevo que excede a la teoría existente. Es decir, cuando los desarrollos teóricos no dan cuenta en su totalidad de nuevos fenómenos sociales o su aparición no fue predicha a partir de la teoría conocida y su impacto en la sociedad es suficientemente relevante como para convertirlo en objeto de observación. Un ejemplo: el estudio de nuevos movimientos sociales a través de alguna organización civil específica.
2) La percepción de un caso, tradicionalmente enmarcado en alguna teoría existente, que presenta especificidades que ésta no alcanza a explicar. Se trata de un caso extremo o desviado en relación con una teoría, que necesita una explicación *ad hoc*.

3) A partir del interés por estudiar un fenómeno general, se selecciona un caso de observación por sus características paradigmáticas o ejemplificadoras. La intención puede ser proveer de base empírica a una teoría o desarrollar un estudio crítico tendiente a su reformulación. En esta última perspectiva el caso se enmarca en la lógica del experimento crucial.
4) Las dificultades (económicas o de accesibilidad) para llevar a cabo un estudio amplio algunas veces desembocan en la elección de uno o algunos casos como una etapa previa a futuras investigaciones más abarcativas.

14.3. Algunas tipologías de estudio de caso

La falta de consenso sobre la caracterización metodológica de esta estrategia investigativa ha dado lugar a la construcción de diversas tipologías de EC, desde diferentes enfoques:

A) Robert Stake (1994) clasifica los EC, a partir del propósito del investigador, en: intrínseco, instrumental y colectivo.

EC INTRÍNSECO: El caso en sí mismo es el que despierta el interés por su estudio. La investigación se orienta a una descripción densa que relate la especificidad de ese caso, sus relaciones y sus particularidades. No existe ningún interés que trascienda la preocupación por el caso en sí. Un ejemplo de estudio intrínseco es la investigación biográfica donde el interés está centrado en la vida de un individuo en particular.

EC INSTRUMENTAL: El caso cumple el rol de mediación para la comprensión de un fenómeno que lo trasciende. El propósito de la investigación va más allá del caso; éste es utilizado como instrumento para evidenciar características de algún fenómeno o teoría. El foco de la atención y la comprensión desborda los límites del caso en estudio. El caso puede ser seleccionado como caso ejemplar o típico.

EC COLECTIVO: Consiste en una investigación comparativa del estudio de varios casos. Puede tratarse de un diseño basado en casos similares o diferentes (Przeworski y Teune, 1970) o de replicaciones (Yin, 1993). Lo que convierte a estas investigaciones en EC es que el conjunto de los casos seleccionados no es considerado como una muestra estadística representativa de una población de donde se infieren generalizaciones. Por el contrario, cada caso es estudiado y comprendido en su especificidad para luego proceder a la comparación entre ellos, dando lugar a la comprensión de una clase mayor de casos. En este sentido la búsqueda no se orienta sólo hacia las características que tienen en común sino también hacia sus diferencias.

Los tipos de EC propuestos por Stake no son excluyentes: un caso intrínseco puede convertirse en un caso ejemplar y un caso instrumental puede acoplarse en un diseño colectivo.

B) Robert Yin (1993) elabora dos tipologías basadas en diferentes criterios. Por un lado distingue, según el número de casos seleccionados, entre el EC único que concentra la atención en un solo caso y el EC múltiple que incluye más de un caso en un mismo estudio, siguiendo la lógica de la replicación. Este último coincide con el EC colectivo de Stake.

El diseño basado en un caso único se justifica cuando se trata de un caso crítico (utilizado para poner a prueba una teoría), de un caso extremo (muy raro o difícil de encontrar) o de un caso revelador (fenómeno de difícil acceso).

En segundo lugar distingue entre el EC holístico y el incrustado (*embedded*). Esta distinción se basa en el número de unidades de análisis que incluye el diseño de la investigación. Se trata de un diseño holístico cuando se trabaja con una sola unidad de análisis (el caso), mientras que en el diseño incrustado se presta atención a subunidades que permiten análisis más amplios para profundizan la comprensión del caso (por ejemplo, se puede estudiar una organización a través de sus reglas y accionar, pero también prestar atención a los individuos que la conforman).

El eje único/múltiple es cruzado por un lado con el eje holístico / incrustado (*embedded*), dando lugar a cuatro tipos de diseño.

Tabla 14.1: *Tipos básicos de diseño para estudios de caso según Yin*

	CASO ÚNICO	CASOS MÚLTIPLES
holístico (una única unidad de análisis)	tipo 1	tipo 3
incrustado (múltiples unidades de análisis)	tipo 2	tipo 4

Fuente: Yin (1993: 39).

Por otro lado, Yin combina el eje único/múltiple con una clasificación de EC basada en el objetivo de la investigación: exploratorio, descriptivo y explicativo, generando en este caso seis modelos.

Tabla 14.2: *Segunda tipología de Yin*

	ÚNICO	MÚLTIPLES
exploratorio	tipo I	tipo IV
descriptivo	tipo II	tipo V
explicativo	tipo III	tipo VI

EC EXPLORATORIO: está orientado a definir preguntas e hipótesis para estudios posteriores más amplios. Se ubica en las primeras etapas de la investigación y muchas veces se utiliza como prueba piloto a partir de la cual se desarrollan marcos conceptuales y operacionales.

EC DESCRIPTIVO: provee una descripción densa del caso contextualmente situado.

EC EXPLICATIVO: a través de este diseño se intenta establecer relaciones de asociación entre las variables bajo estudio en el caso. Se orientan a ser estudios causales.

C) Arend Lijphart (1971) considera que la discusión del método comparativo no está completa si no se toma en consideración el método de EC. Diferencia a los métodos sobre la base del número de casos: caracteriza el método estadístico por basarse en muchos casos, el comparativo por aplicarse a pocos casos y el EC por aplicarse sólo a uno. Teniendo en cuenta esta distinción, afirma que en un estudio de caso no es posible inferir una generalización ni refutar generalizaciones establecidas. Sin embargo, le reconoce el mérito de contribuir al proceso de formulación de proposiciones y construcción de teorías a partir de los resultados de un análisis intenso del objeto. En este sentido, enuncia seis tipos de EC que caracteriza como tipos ideales, de modo que un estudio determinado puede adecuarse a más de una de las categorías. Los dos primeros focalizan en el caso en sí mismo (el caso intrínseco de Stake resulta similar en este sentido), y los cuatro últimos se vinculan a la elaboración de teorías:

EC ATEÓRICOS: se trata de investigaciones totalmente descriptivas que resultan útiles como fuente exploratoria de información. Este tipo se asemeja en parte al diseño "exploratorio basado en un caso" de Yin con algunas características del caso intrínseco de Stake. Su especificidad está dada porque Lijphart lo despoja de toda relación con la teoría. Sus resultados podrían contribuir indirectamente al proceso de construcción de teoría sólo en la medida en que fueran interpretados en el marco de un análisis teórico posterior.

EC INTERPRETATIVOS: el interés sigue puesto en el caso más que en una teoría general, pero a diferencia de los anteriores utilizan proposiciones de la teoría establecida en forma explícita. Estas generalizaciones se orientan a iluminar la comprensión del caso, en consecuencia este tipo de EC también carece de valor para la construcción teórica.

EC GENERADORES DE HIPÓTESIS: intentan comprobar hipótesis iniciales vagas y a partir de allí desarrollar generalizaciones teóricas en áreas poco exploradas. Tienen un alto valor teórico.

EC QUE CONFIRMAN UNA TEORÍA: se orientan a la puesta a prueba de proposiciones e hipótesis. Se trata de ejemplos corroborativos.

EC QUE DEBILITAN UNA TEORÍA: igual que el anterior, se orienta a la puesta a prueba de proposiciones, pero en este tipo éstas podrían no ser confirmadas por el caso en cuestión. Según Lijphart, a partir de un solo caso no es posible refutar una teoría pero sí debilitarla.

EC DESVIADOS: surgen con la intención de comprender o explicar casos que se desvían de las teorías generales. Tienen alto valor teórico porque, si bien debilitan las proposiciones originales, dan lugar a nuevas proposiciones que pueden tener mayor capacidad heurística.

Según Lijphart (1971) los estudios generadores de teorías y los de casos desviados son los que más contribuyen a la construcción de teoría.

D) Giovanni Sartori (1990) retoma las categorías enunciadas por Lijphart pero toma en cuenta sólo aquellas que se vinculan a la teoría y las denomina estudio de caso heurístico, crucial y desviado:

EC HEURÍSTICO: son los que resultan útiles para la generación de hipótesis
EC CRUCIAL: están orientados a fortalecer o debilitar una teoría
EC DESVIADO: son los que se desvían de una proposición general o teoría establecida.

E) Bent Flyvbjerg (2004) formula una tipología de cuatro tipos de EC que no considera excluyentes entre sí:

EC EXTREMO O DESVIADO: se orienta a los casos inusuales dentro del marco de una teoría. Es semejante al tipo de caso desviado de Lijphart (1971).

LA MÁXIMA VARIACIÓN DE CASOS: es un diseño de casos múltiples seleccionados por sus disimilitudes. Este esquema es semejante a "los casos más diferentes" de Przeworski y Teune (1970).

EC CRÍTICO: los casos críticos permiten inferir conclusiones para el tipo en su conjunto. La criticidad del caso está vinculada a la presencia de todas las propiedades que se intenta estudiar, lo que permite afirmar que las conclusiones válidas para este caso también lo serán para todos los casos del mismo tipo. Es semejante en su función a los EC confirmadores y debilitadores de teorías de Lijphart (1971).

EC PARADIGMÁTICO: son casos que resaltan características más generales de la sociedad. Puede equipararse al que otros autores denominan caso ejemplar o típico y al tipo instrumental de Stake (1994).

La mayoría de las tipologías presentadas podrían reducirse a cuatro tipos básicos definidos a partir de dos ejes: a) el número de casos en estudio, y b) la intención prevaleciente en la selección de los casos —a partir de una lógica empiricista inductiva centrada en el caso mismo o focalizando en el valor heurístico del caso en vinculación con la teoría.

Tabla 14.3: *Tipología de estudio de caso/s*

	NÚMERO DE CASOS	
TIPO DE DISEÑO	UNO	MÁS DE UNO
holístico (centrado en el caso)	intrínseco (Stake) ateórico (Lijphart) interpretativo (Lijphart)	
heurístico (vinculado a la teoría)	instrumental (Stake) desviado (Lijphart, Sartori, Flyvbjerg) crítico (Flyvbjerg) paradigmático (Flyvbjerg) generador de hipótesis/heurístico (Lijphart, Sartori) confirmador/debilitador de teoría// ejemplificador/ típico/ promedio//crucial (Lijphart, Sartori)	colectivo (Stake) múltiple (Yin)

Basta una mirada a la tabla para concluir que los autores han prestado mayor atención a la categorización de los estudios de caso único vinculados con la construcción de teoría. Para los comparativistas (Lijphart, Eckstein, Sartori, Bartolini), los estudios basados en más de un caso —que implican necesariamente la comparación— se enmarcan en el método comparativo y no en el denominado EC. Por otro lado, en general, se otorga más reconocimiento a los EC con valor heurístico, en la medida que son capaces de proveer soporte empírico, elementos conceptuales y núcleos de sentido para la comprensión de fenómenos que los trascienden.

Sin embargo, la función del EC en tanto control de una teoría no está exenta de debate. Collier (1991: 74) destaca "la tensión entre investigaciones basadas en muchos casos y pocas variables [método estadístico] e investigaciones orientadas a analizar pocos casos en toda su complejidad [EC]".

La diversidad de tipologías (las presentadas aquí, si bien son las más utilizadas, no agotan las existentes) pone de manifiesto la importancia de caracterizar cada EC, tanto en sus criterios definitorios como en los objetivos que motivan su aplicación, a fin de esclarecer su alcance y relevancia. Éste es un punto importante a ser tenido en cuenta por los estudiantes de posgrado en sus propuestas de tesis basadas en estudios de caso, que resultan numerosas dada la dificultad que tienen para acceder a recursos que permitan diseñar estudios con mayor cantidad de casos y por un criterio pragmático orientado a la selección de un objeto fácil de abordar.

14.4. La selección de los casos

Cuando el caso no viene definido desde fuera de la investigación (como en el ejemplo terapéutico) o cuando no se trata de un caso intrínseco (donde es el propio objeto el que despierta el interés del investigador) surge el problema de su selección.

En primer lugar es necesario definir la unidad de análisis (individuos, ciudades, políticas públicas, etcétera), para luego seleccionar dentro del conjunto de unidades de análisis el caso (el caso Juanito, el caso de la ciudad de Buenos Aires, la Ley Nacional de Educación). La selección de la unidad de análisis está condicionada por la perspectiva teórica y los objetivos de la investigación, y una vez que ésta ha sido determinada, ella condiciona la estrategia de la recolección de datos.

Si todas las unidades de análisis fueran idénticas en las propiedades relevantes para los objetivos de la investigación, cualquiera de ellas podría convertirse en "el caso". Pero, dada su diversidad, es necesario definir los criterios que convertirán a una de las unidades de análisis en el caso a estudiar.

El proceso de selección entre las unidades es diferente para cada uno de los tipos de EC. Si se trata de un EC intrínseco, éste es seleccionado sólo a partir de sus propias características y del interés que éstas despiertan, mientras que para los instrumentales y colectivos la selección se produce a partir de ciertos criterios establecidos por el investigador, que los convierte en foco de atención. De-

pendiendo de los objetivos propuestos se puede elegir, por ejemplo, estudiar un caso típico (ejemplar) o un caso desviado (Feagin, Orum y Sjoberg, 1991). El caso típico o promedio no es necesariamente el que revela mayor información; los casos atípicos o extremos algunas veces proveen más y mejores elementos para comprender los fenómenos (Flyvbjerg, 2006).

La selección del caso, entonces, nunca se realiza por azar como en los métodos estadísticos, porque se parte de la no equivalencia entre los casos, que se destaca a través de la propuesta orientada a la comprensión de la especificidad de cada caso. El proceso de selección está basado en criterios teóricos, en experiencias de observación y en las expectativas depositadas en la unidad seleccionada, en términos de su potencialidad para proveer una base empírica relevante para la interpretación y comprensión del fenómeno estudiado.

La dificultad en la elección del caso reside en su íntima relación con el contexto, que hace necesario buscar los criterios de distinción que permitirán individualizar el objeto. El primer problema, entonces, en el proceso de selección del caso es su definición.

La importancia de definir los límites (acotar el caso) estriba en que ellos establecen la frontera entre lo que la investigación incluye y lo que queda fuera de ésta. Esta decisión fundante del caso es tarea del investigador, quien, a través de ella, construye su caso de estudio.

Ragin y Becker (1992) plantean que los casos pueden ser vistos como unidades empíricas o como construcciones teóricas. Es decir, desde una perspectiva empiricista que afirme la transparencia de los hechos, se considera que tienen una existencia real; mientras que desde un enfoque constructivista donde la realidad es el resultado de la interpretación, son producto de convenciones o consensos teóricos.

Stake (1995) enuncia dos criterios de selección: maximizar lo que potencialmente podemos aprender del caso y su mayor accesibilidad. Es decir, es conveniente elegir aquellos casos que parecen ofrecer mayores oportunidades para aprender y aquellos más fáciles de abordar y donde se espera una mayor receptividad.

Es importante destacar que el proceso de selección y sus criterios no se agotan en la identificación del caso; el investigador también debe establecer los niveles internos de análisis —es decir hasta dónde se abordarán las complejidades del caso— teniendo en cuenta que él opera dentro de una diversidad de contextos. En este sentido, si bien el caso intrínseco (o el que viene dado) no se elige dentro de un conjunto de unidades de análisis, en el proceso de su estudio también se lleva a cabo un conjunto complejo de elecciones.

Dado que la selección no es azarosa y que la participación de la subjetividad es alta, uno de los riesgos es la selección desviada, en particular, cuando se trata de estudios de casos para confirmar o debilitar generalizaciones. De allí la importancia de hacer explícitos los criterios de selección del caso y su relación con la teoría.

CAPÍTULO 15
EL ANÁLISIS BIVARIABLE

En este capítulo se ilustra la forma más básica de analizar las relaciones entre las propiedades que ya se transformaron en variables en la matriz de datos. Esta forma se denomina análisis bivariable porque sólo se consideran relaciones entre dos variables; a pesar de su carácter elemental, hay que franquear esta puerta para acceder a formas más complejas de análisis.

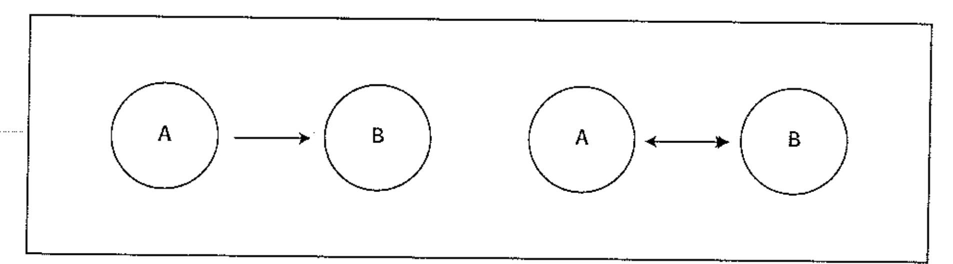
Por lo demás, si se elige agudamente la pareja de variables que se relacionan, a veces este análisis sencillo puede echar luz sobre un análisis sofisticado cuyos resultados sean difíciles de interpretar.

15.1. Los modelos

Un modelo es la representación gráfica de una hipótesis acerca de la relación que se quiere investigar. Sus elementos son sencillos: círculos para representar las variables y flechas para representar las relaciones entre ellas.

La figura 15.1 presenta las formas que puede tomar un modelo bivariable. Si nuestra hipótesis es que la variable A influye en la variable B sin ser influida por B, hay una flecha que parte de A y llega a B. En este caso, A se denomina variable INDEPENDIENTE y B se denomina variable DEPENDIENTE. Si en cambio la hipótesis es que las dos variables se influyen mutuamente, se representa con una flecha con dos puntas (a veces se pone una simple línea en vez de este tipo de flecha).

Figura 15.1: *Dos modelos bivariables*



A menudo la relación A←→B es denominada simétrica. El sentido griego del término conlleva la idea de que las dos variables se influyen "en la misma medida", es decir con la misma fuerza. Pero este tipo de relación pasa muy raramente en la realidad, y la mayoría de las relaciones son bidireccionales asimétricas (la influencia de A en B es más fuerte que la de B en A).

Por otro lado, no existen técnicas que puedan discriminar entre relaciones simétricas y asimétricas, y eso explica el uso del término 'simétrico' en lugar del más correcto 'bidireccional'. El término 'asimétrica' se usa a menudo para la relación A→B, que debería llamarse 'unidireccional'.

Cabe subrayar de inmediato que sólo en el marco de un experimento (véase apartado 1.1) se pueden efectivamente controlar nuestras hipótesis acerca de la dirección de una influencia entre dos variables. Pero los límites del modelo experimental cuando se aplica a problemas humanos y sociales (véase apartado 1.2) nos obligan a valernos del método de la asociación. Con este método, la dirección que tiene la influencia entre dos variables puede ser imaginada basándose en nuestro conocimiento común, pero no puede ser averiguada impersonalmente, sólo basándose en las técnicas de análisis (véanse los apartados 1.3 y 1.4). Esto es así porque sólo se pueden registrar los estados de los objetos/sujetos en las diversas propiedades, pero no se pueden alterar los estados en una propiedad para ver el efecto de esta manipulación en otras.

Dada esta situación, si se le pide a la computadora analizar la relación entre A y B, ella puede cuantificar la intensidad de esta asociación a través de coeficientes[1] adecuados y (si las variables son ordinales o cardinales) establecer si la asociación es positiva o negativa, pero nada más. En cuanto a la dirección de esta influencia, la técnica tiene que aceptar la que impone nuestro modelo. Es el modelo —es decir, el juicio del investigador— el que conlleva recurrir a técnicas y coeficientes que asumen que la relación es unidireccional o que es bidireccional.

En otras palabras: si se pide a la computadora que nos diga con qué intensidad el ingreso de cada individuo influye en su edad, la computadora va dar su (absurda) respuesta: por ejemplo, que cada cien pesos más de ingreso mensual se produce un envejecimiento medio de un año en quien lo cobra. Al analizar sólo vectores de cifras no interpretadas, la computadora no puede hacer otra cosa. Le corresponde al investigador elegir modelos que asumen entre las variables las mismas direcciones de influencia que existen entre las propiedades en la realidad. De todas maneras, él no puede declarar que los resultados de su

[1] En matemática se denomina 'coeficiente' un número que multiplica otro número o expresión numérica. En análisis de datos el término se usa para indicar la cifra que cuantifica la intensidad de una asociación entre dos o más variables. La gran mayoría de los coeficientes que se usan son "estandarizados", es decir que su valor puede variar entre −1 (una asociación negativa de la máxima intensidad posible) y +1 (una asociación positiva de la máxima intensidad posible). El valor 0 indica independencia, es decir ausencia de asociación entre las variables. Para indicar que un coeficiente es estandarizado, para sus valores decimales se usa convencionalmente la notación sin cero y con punto en lugar de la coma (por ejemplo: .63). En cambio, los coeficientes no estandarizados (como *b*, el coeficiente de regresión: véase el apartado 15.6) se indican con la notación con cero y coma (por ejemplo: 0,63).

análisis confirman su juicio acerca de la dirección: si un resultado no puede desmentir una hipótesis, tampoco puede confirmarla.

Adoptar un modelo A←→B o un modelo A→B no implica de ninguna manera la convicción de que las propiedades correspondientes a las variables A y B no son influidas en la realidad por otras propiedades cuyas variables no fueron tenidas en cuenta en aquel modelo. Por el contrario, un modelo bivariable es —como ya se dijo— la puerta de acceso a modelos más complejos, que tratan de cuantificar al menos un segmento de la tupida red de relaciones que existe en la realidad. Los límites propios de las técnicas no permiten a ningún modelo aproximarse lo suficiente a la realidad; esto hay que tenerlo siempre presente cuando se analicen los resultados de nuestros modelos.

En matemática se acostumbra a imaginar situaciones en que los valores de una variable dependen exclusivamente de los valores de otra, tal como las representa la función $Y = f(X)$. Pero incluso en las ciencias físicas y biológicas se sabe que este modelo no es realista: en cada situación específica el valor de Y es influido por otras propiedades además de X. Si estas influencias son accidentales y dispersas, se pueden considerar globalmente como "error", y el modelo se convierte en $Y_i = f(X_i) + e_i$. Esto significa que en cada caso (subíndice i) la variable Y asume un valor que depende del valor de la variable X en ese caso y del complejo de perturbaciones (el "error") que actúan en él. Pero si algunas propiedades diferentes de X influyen sistemáticamente en Y, ellas tienen que ser insertadas en el modelo, que entonces deja de ser bivariable.

En las ciencias sociales, un modelo que vincula una variable sólo a la influencia de otra es demasiado lejano de lo que pasa en la realidad como para ser útil en nuestro análisis. Por otro lado, sólo un conocimiento profundo de las técnicas bivariables permite manejar las técnicas más complejas sin cometer burdos errores metodológicos. En este espíritu se presentarán en lo que sigue las principales técnicas de análisis bivariable.

15.2. La relación entre dos dicotomías

Con el término, de origen griego, dicotomía se denomina una variable que sólo tiene dos categorías. La herramienta con la que se analiza la relación entre dos dicotomías se denomina TABLA BIVARIABLE (o de contingencia). En las tablas 15.1, 15.2 y 15.3 se presenta la relación entre dos dicotomías (el sexo y el haber votado en las últimas elecciones) en una hipotética muestra de adultos. Esta tabla puede tomar tres aspectos según los intereses de quien la construye.

En la tabla 15.1A se presentan las FRECUENCIAS EMPÍRICAS: en la celda de arriba a la izquierda, la cifra 226 significa que 226 mujeres declararon haber votado. En la celda a su derecha se ve que 84 mujeres declararon no haber votado. Y así para los varones en las celdas de la segunda fila. Las cifras al margen de la tabla (que se denominan MARGINALES) nos dicen que las mujeres de la muestra son 310 y los varones son 295, sobre un total de 605, y que 468 votaron y 137 no votaron.

Tabla 15.1: *Frecuencias y porcentajes*

A: FRECUENCIAS EMPÍRICAS				B: PORCENTAJES DE FILA				C: PORCENTAJES DE COLUMNA			
	SÍ	NO	TOT.		SÍ	NO	TOT.		SÍ	NO	TOT.
M	226	84	310	M	73	27	100	M	48	61	51
V	242	53	295	V	82	18	100	V	52	39	49
	468	137	605		77	23	100		100	100	100

Si queremos saber en qué porcentaje las mujeres declaran haber votado, convertimos en cien el total de las mujeres y expresamos 226 como porcentaje de 310. Lo mismo para los varones, y para el total de la muestra. Siempre que los marginales de fila sean 100 como en la tabla 15.1B, la tabla presenta los PORCENTAJES DE FILA. Por otro lado, si queremos saber en qué porcentaje los que declaran haber votado son mujeres (o varones), convertimos en cien los marginales de columna, y expresamos 226 y 242 como porcentajes de 468. Siempre que los marginales de columna sean 100 como en la tabla 15.1C, la tabla presenta los PORCENTAJES DE COLUMNA.

La tabla 15.1B es la más clara en mostrar que el género tiene una cierta influencia sobre el voto (más exactamente, sobre el voto declarado). En el total de la muestra, el 77% declaró haber votado. Si los varones hubieran votado en el mismo porcentaje, serían[2] 228 en vez de 242. Este 228 es la frecuencia teórica de la celda en cuestión. Para calcular rápidamente la FRECUENCIA TEÓRICA en una celda se multiplica el marginal de la fila por el marginal de la columna que se cruzan en aquella celda; ese producto se divide por el total de los casos en la tabla.

La diferencia entre la frecuencia empírica y la frecuencia teórica se denomina DESVIACIÓN. Si en una celda esta diferencia es positiva (la frecuencia empírica es más alta que la teórica) se dice que hay ATRACCIÓN entre las dos categorías que forman la celda. Si esta diferencia es negativa se dice que hay REPULSIÓN entre aquellas categorías. Si las diferencias entre frecuencias empíricas y frecuencias teóricas son pequeñas, se concluye que las dos variables son independientes (no tienen relación entre sí).

Naturalmente, por necesidad matemática desviaciones positivas y desviaciones negativas tienen que balancearse: la suma algebraica de las desviaciones es necesariamente cero. Esto no sólo en toda la tabla, sino también en cada fila y en cada columna.

La tabla 15.2A muestra las desviaciones entre las frecuencias empíricas presentadas en la tabla 15.1A, cuya celda izquierda superior dice que 226 mujeres declararon haber votado. La frecuencia teórica en aquella celda es (310 * 468 / 605 =) 239,8. La desviación es (redondeando a 240) –14. La celda derecha superior dice que 84 mujeres declararon no haber votado. La frecuencia teórica en aquella celda es (310 * 137 / 605 =) 70,2. La desviación es (redondeando a 70) +14. La celda izquierda inferior dice que 242 varones declararon haber votado. La frecuencia teórica en aquella celda es (295 * 468

[2] El cociente exacto es 228,19834711... Pero al no poder fraccionar individuos se redondea.

/ 605 =) 228,2. La desviación es (redondeando a 228) +14. La celda derecha inferior dice que 53 varones declararon no haber votado. La frecuencia teórica en aquella celda es (295 * 137 / 605 =) 66,8. La desviación es (redondeando a 67) –14.

Las desviaciones tienen el mismo valor absoluto en todas las celdas: lo que cambia es el signo. Este resultado no depende de una tabla particular: siendo una necesidad matemática, es válida para todas las tablas —como lo muestra la tabla 15.2B.

Tabla 15.2: *Cómo se balancean las desviaciones*

A: EN UNA TABLA PARTICULAR				B: EN CUALQUIER TABLA			
	SÍ	NO	TOT.		Y_a	Y_b	
M	–14	+14	0	X_a	–n	+n	0
V	+14	–14	0	X_b	+n	–n	0
	0	0	0		0	0	0

Esta necesidad matemática tiene una consecuencia indirecta pero inevitable: si dos dicotomías tienen frecuencias desequilibradas, no merece la pena relacionarlas, porque una (eventual) relación no puede aflorar con bastante nitidez.

Veamos por qué, siguiendo con nuestro ejemplo de la relación entre género y voto. Pero imaginemos que en la muestra casi todas son mujeres y casi todas declararon haber votado, como en la tabla 15.3A.

Tabla 15.3: *Frecuencias desequilibradas*

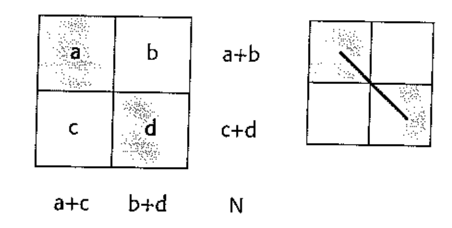
A: FRECUENCIAS EMPÍRICAS				B: FRECUENCIAS TEÓRICAS				C: DESVIACIONES			
	SÍ	NO	TOT.		SÍ	NO	TOT.		SÍ	NO	TOT.
M	40	0	40	M	36	4	40	M	+4	–4	0
V	5	5	10	V	9	1	10	V	–4	+4	0
	45	5	50		45	5	50		0	0	0

En la tabla 15.3B se presentan las frecuencias teóricas y en la 15.3C las desviaciones. En la celda derecha inferior la frecuencia teórica es (5 * 10 / 50 =) 1 y la frecuencia empírica es la máxima posible: ya que los que no votaron son 5, los varones que no votaron no pueden ser más que 5. De eso se sigue que la desviación de (5 – 1 =) +4 es la máxima posible en (todas las celdas de) la tabla. Sin embargo, a pesar de ser la más alta posible, la desviación no es bastante alta como para llegar a la conclusión de que existe una asociación entre el género y la propensión al voto.

Un desequilibrio entre las frecuencias de una de las dicotomías (o de ambas) que se relacionan conlleva también un sesgo en los valores que cobran los coeficientes con los que se cuantifica la intensidad de la asociación entre dos dicotomías.

Para explicar cómo y por qué pasa esto, hay que introducir la convención simbólica que siguen todos los coeficientes. Las frecuencias en las cuatro celdas se indican con las primeras cuatro letras del alfabeto, en el orden que muestra la tabla 15.4. Por consiguiente, los marginales y el total se indican en la forma que presenta la misma tabla. La diagonal que une las celdas *a* y *d* se llama diagonal principal, y se dice que entre las dicotomías hay una asociación si las celdas en esa diagonal presentan desviaciones positivas respecto de las frecuencias teóricas. Por el contrario, si las desviaciones positivas se presentan en la otra diagonal, se habla de una relación negativa.

Tabla 15.4: *Celdas, marginales y diagonal*



Casi todos los coeficientes que fueron imaginados para cuantificar la asociación entre dos dicotomías aplican esta convención, basándose en el llamado PRODUCTO CRUZADO (ad, bc) que constituye el numerador de la fórmula. Si el producto entre las frecuencias de la diagonal principal supera el producto entre las frecuencias de la otra diagonal, el coeficiente será positivo.

Para entender por qué el desequilibrio de las frecuencias de las dicotomías (que se convierten en marginales de la tabla) conlleva un sesgo en los coeficientes, necesitamos destacar una propiedad de cada producto. Cuanto más desequilibrados sean los multiplicandos, tanto menor será el producto. Esta simple noción matemática se puede comprobar fácilmente considerando una secuencia de productos: $5 * 5 = 25; 6 * 4 = 24; 7 * 3 = 21; 8 * 2 = 16; 9 * 1 = 9$. Ya que todos los coeficientes tienen productos en el numerador y en el denominador, se entiende fácilmente que si las frecuencias en las celdas son desequilibradas por efecto de un desequilibrio en los marginales, eso va a producir un sesgo en el valor numérico del coeficiente.

Unos ejemplos mostrarán cómo funciona este mecanismo. Ya que la tabla 15.3A presentaba un caso de relación entre dos dicotomías desequilibradas, se puede empezar con ella, aplicando a la tabla los coeficientes dicotómicos más usados,[3] cuyas fórmulas se reportan a continuación.

[3] Hay muchos coeficientes más, pero todos presentan el mismo defecto, ya que incluyen multiplicaciones en sus fórmulas. Entre los coeficientes que se presentan en este libro, Q fue propuesto por Yule en honor al estadístico Quetelet; φ (Fi) es una letra griega correspondiente a una f; τ (tau) es una letra griega correspondiente a una t. Ya que hay muchos coeficientes diferentes, se usan afijos para distinguirlos.

$$Q = \frac{ad - bc}{ad + bc}$$

Poniendo en la fórmula las frecuencias de la tabla 15.3A, tenemos:

$$Q = \frac{40 * 5 - 0 * 5}{40 * 5 + 0 * 5} = \frac{200}{200} = 1$$

Por lo tanto, el valor de Q nos dice que la asociación es la máxima posible, pero en realidad sólo sería tal si ninguno de los varones votara.

Se puede razonablemente pensar que este grave sesgo sólo se produce porque hay una frecuencia igual a cero. Veamos qué pasa si aplicamos la misma fórmula a la tabla 15.5, donde no hay frecuencias nulas:

Tabla 15.5: *Otras frecuencias desequilibradas*

	SÍ	NO	TOT.
M	45	2	47
V	4	3	7
	49	5	54

$$Q = \frac{45 * 3 - 2 * 4}{45 * 3 + 2 * 4} = \frac{135 - 8}{135 + 8} = .89$$

También en este caso Q sobrevalora claramente la fuerza de la asociación presente en la tabla.

Veamos ahora qué pasa con el coeficiente φ (Fi), cuya fórmula es:

$$\varphi = \frac{ad - bc}{\sqrt{(a + b)(c + d)(a + c)(b + d)}}$$

Para la tabla 15.3A, φ cobra el valor

$$\varphi = \frac{40 * 5 - 0 * 5}{\sqrt{40 * 10 * 45 * 5}} = \frac{200}{300} = .67$$

Para la tabla 15.5, φ cobra el valor

$$\varphi = \frac{45 * 3 - 2 * 4}{\sqrt{47 * 7 * 49 * 5}} = \frac{127}{283{,}91} = .45$$

El nivel del coeficiente indica que φ también sobrevalora, sin embargo significativamente menos que Q.

El tercer coeficiente que aplicamos es τ_c (Tau$_c$), cuya fórmula para dos dicotomías es:

$$\tau_c = \frac{4 * (ad - bc)}{N^2}$$

Para la tabla 15.3A, τ_C cobra el valor

$$\tau_c = \frac{4 * 200}{2.500} = .32$$

Para la tabla 15.5, τ_C cobra el valor

$$\tau_c = \frac{4 * 127}{54^2} = \frac{508}{2.916} = .17$$

En ambos casos, los valores que cobra τ_C son mucho más aceptables que los otros. Pero no es siempre así. Por ejemplo, en la tabla 15.6A, τ_C cobra el valor (4 * 300) / 1600 = 1200 / 1600 = .75. Pero la asociación parece perfecta, como en efecto la estiman Q y φ.

Tabla 15.6: *Ejemplos de asociaciones perfectas y casi perfectas*

A: PERFECTA	SÍ	NO	TOT.	B: CASI PERFECTA	SÍ	NO	TOT.
M	30	0	30	M	20	1	21
V	0	10	10	V	2	77	79
	30	10	40		22	78	100

En la tabla 15.6B, τ_C cobra el valor [4 * (1.540 – 2) / 10.000 =] .615. Pero la asociación parece más estricta, como en efecto la estima φ (.91), mientras que Q todavía la sobrevalora (.997).

No siempre τ_C subestima como en estos casos; a veces sobrevalora, como en la tabla 15.7, donde cobra valor [4 * (696 - 27) / 81^2 =] .41. Pero la asociación entre las dos dicotomías parece más débil.

Tabla 15.7: *Un caso de asociación menos estricta de lo que parece*

	SÍ	NO	TOT.
M	24	1	25
V	27	29	56
	51	30	81

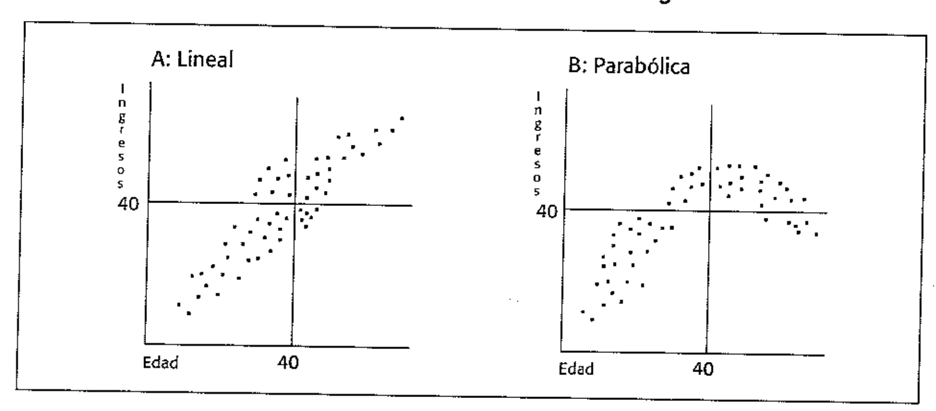
No hay un coeficiente que esté exento de estos riesgos de sesgo cuando los marginales de las dicotomías son desequilibrados. El hecho de que el sesgo se manifieste, y su gravedad, dependen de dónde estén los desequilibrios en la tabla en relación con dónde están los productos en la fórmula del coeficiente. El lector puede encontrar útil la tabla 15.8, en la que se presentan varias formas de marginales de las dicotomías, y las consecuencias en los tres coeficientes que hemos analizado.

Tabla 15.8: *Desequilibrio de los marginales y sesgo de los coeficientes*

LOS MARGINALES SON DESEQUILIBRADOS Y...	Q	φ	τ_c
una celda está (casi) vacía	sobrestima muchísimo	sobrestima mucho	sobrestima mucho
dos celdas (en diagonal) están (casi) vacías	es bueno	es bueno	puede subestimar mucho
tres celdas están (casi) vacías	sobrestima muchísimo	sobrestima mucho	es aceptable

Antes de terminar este apartado, cabe destacar las malas consecuencias de una práctica muy difundida entre los que analizan los datos de las ciencias sociales. Hablamos de la tendencia a convertir en burdas dicotomías propiedades que tienen más que dos categorías significativas, e incluso que son continuas, como el ingreso y la edad. Esta práctica no sólo conlleva una pérdida de información sin algún motivo, sino que también puede entrañar representaciones tergiversadas de la realidad. Para comprobarlo, consideramos los dos diagramas en la figura 15.2.

Figura 15.2: *Relaciones entre edad e ingreso*



La figura 15.2B representa de forma realista la relación entre edad e ingreso: el ingreso (en la ordenada) de un individuo tiende a crecer con el crecimiento de su edad hasta cuando él se retire, a los 60 años; luego tiende a reducirse. En cambio, la figura 15.2A representa una relación lineal entre edad e ingreso: el ingreso tiende a crecer indefinidamente al crecer la edad. Si tratamos las propiedades continuas como edad e ingreso en forma adecuada y las medimos, la diferencia entre las dos situaciones es evidente. Pero si las convertimos en burdas dicotomías y registramos los sujetos como mayores (por encima de 40 años) o como jóvenes (40 años o menos) y como ricos (por encima de 40 mil pesos de ingreso anual) o pobres (40 mil pesos o menos) ambas situaciones quedan representadas en la misma tabla 15.9. Como consecuencia de la dicotomización, la diferencia entre las dos situaciones se pierde completamente.

Tabla 15.9: *Dos propiedades continuas convertidas en dicotomías*

	JÓVENES	MAYORES
ricos	8	18
pobres	25	6

15.3. La relación entre dos variables categoriales

La herramienta con la que representamos la relación entre dos variables categoriales es la misma que para la relación entre dos dicotomías: la tabla bivariable. Es válido también mucho de lo que dijimos en el apartado precedente respecto de marginales, frecuencias empíricas, teóricas y desviaciones, porcentajes de fila y de columna, atracción y repulsión entre categorías. Por el contrario, la diagonal no tiene ninguna relevancia particular, e incluso no existe si las dos variables tienen un número diferente de categorías.

El criterio a seguir para analizar una tabla de esta naturaleza se ilustrará comentando la tabla 15.10, en la que se presenta la relación entre la ocupación y el partido preferido en una hipotética muestra de electores de un país europeo.

Para mostrar cómo se analiza una tabla bivariable, empecemos con las dos primeras celdas en la segunda fila, en las que se comprueba que 12 obreros han declarado haber votado a los comunistas y 48 a los socialistas. Las frecuencias teóricas en estas celdas son (123 * 48 / 1.410 =) 4,19 y (123 * 397 / 1.410 =) 34,6. Se concluye que hay atracción entre tener la ocupación de obrero y votar a los socialistas (desviación de +13,4) y a los comunistas (desviación de casi +8).

Naturalmente, la conclusión que concierne a la celda de los obreros socialistas es más confiable que la otra, al estar basada en una frecuencia teórica más alta, lo que tiende a descartar la posibilidad de que sea un efecto del azar (error muestral: véase el apartado 6.3). Por el contrario, la atracción entre amas de casa y voto al partido de la izquierda cristiana (frecuencia empírica de 5, frecuencia teórica de 2,5, desviación de +2,5) no es bastante confiable por estar basada en un pequeño número de casos.

Veamos ahora la celda de los 17 comerciantes que declararon haber votado a los socialistas. La frecuencia teórica en esa celda es (74 * 397 / 1.410 =) 20,84 y la desviación es −3,84, suficiente para concluir que hay repulsión entre las dos categorías (ser comerciante y votar a los socialistas). Esta conclusión es confiable por estar basada en una frecuencia teórica bastante alta. A la inversa, la repulsión que se comprueba entre los campesinos y el voto al partido liberal (frecuencia empírica de 2, frecuencia teórica de 5,1, desviación −3,1) parece fuerte pero no es tan confiable al estar basada en una frecuencia teórica baja.

Por fin veamos la celda de los 37 profesionales que declararon haber votado al partido conservador. La frecuencia teórica es casi igual a la empírica, siendo (186 * 289 / 1.410 =) 38,1. La conclusión de que entre las dos categorías no existe ni atracción ni repulsión es confiable al estar basada en una alta frecuencia teórica.

Tabla 15.10: *La relación entre dos variables categoriales*

	COMUNISTA	SOCIALISTA	IZQ. CRIST.	LIBERAL	CONSERV.	NACIONAL.	NR	TOTAL
campesino	–	6	3	2	13	–	18	42
obrero	12	48	3	6	9	4	41	123
policía y militar	–	4	–	1	15	7	24	51
oficinista	7	50	5	8	35	5	52	162
artesano	3	24	3	11	15	5	26	87
maestro y profesor sec.	5	21	3	12	10	3	8	62
técnico	1	26	3	9	19	1	25	84
funcionario	–	16	–	8	13	2	8	47
concesionario	–	13	–	10	14	2	19	58
comerciante	1	17	2	11	18	4	21	74
profesional	7	55	4	31	39	4	46	186
prof. univers.	3	12	2	6	6	1	1	31
empresario	–	4	–	8	17	4	9	42
ama de casa	–	24	5	11	23	2	33	98
estudiante	5	17	2	11	4	1	15	55
jubilado	4	60	1	27	39	6	71	208
total	48	397	36	172	289	51	417	1.410

En resumidas cuentas: el primer paso para analizar una tabla bivariable es calcular las frecuencias teóricas y las desviaciones en cada celda: no es un gran problema porque todos los programas informáticos de análisis lo hacen por nosotros. El segundo paso es seleccionar las celdas que presentan una frecuencia teórica bastante alta (al menos 10 casos) y por lo tanto permiten sacar conclusiones bastante confiables. El tercer paso es considerar las desviaciones entre frecuencias empíricas y frecuencias teóricas. En las celdas en que la desviación es positiva y no despreciable (al menos el 5 o 10% de la frecuencia teórica misma) se puede comprobar que hay atracción entre las dos categorías que forman la celda. Análogamente para las celdas en las que la desviación es negativa y no despreciable: si la frecuencia teórica fue bastante alta se puede detectar la presencia de un rechazo entre las dos categorías que forman la celda. Si frecuencias empíricas y frecuencias teóricas son iguales o cercanas, no hay ni atracción ni repulsión.

Por curiosidad se puede replicar el mismo análisis con las celdas que hemos descartado al comienzo por tener bajas frecuencias teóricas, pero sin olvidar que nuestras conclusiones no serán confiables. Una decisión más sabia sería analizar las categorías de ambas variables de la tabla para ver si algunas pueden ser agregadas siguiendo las pautas establecidas en el apartado 8.1.

Consideramos ahora nuevamente la tabla 15.10. ¿Podemos decir que atestigua la presencia de relaciones entre ocupación y voto? Seguramente sí, porque muchas celdas muestran atracciones o repulsiones, pese a que algunas no permiten conclusiones confiables. ¿Podemos decir que hay una relación global? Seguramente no, porque cada celda presenta su situación particular. Saber que

hay relaciones entre ocupación y voto tiene escaso interés si no se sabe qué ocupación favorece votar a qué partido. Por consiguiente, no tienen sentido los coeficientes que tratan de cuantificar la fuerza de la relación global entre dos variables categoriales (φ^2, C de Pearson, V de Cramér: véase Blalock 1960, par. XV.3).

Todos estos coeficientes son elaboraciones[4] de un conocido coeficiente global, el χ^2 (chi-cuadrado),[5] que no cuantifica la intensidad de la asociación, sino la improbabilidad de que se forme una tabla dada con datos de una muestra si las dos propiedades son independientes[6] en la población. La fórmula de chi-cuadrado está basada en las desviaciones entre las frecuencias empíricas (observadas) y las frecuencias teóricas (esperadas); en la fórmula se usan las iniciales de los adjetivos ingleses, *observed* y *expected*: ver más adelante. Cada desviación está normalizada dividiéndola por el valor de la frecuencia teórica, por el obvio motivo de que una desviación de 2 es diez veces más importante si deriva de una diferencia 12-10 que si deriva de una diferencia 112-110. Por otro lado, antes de sintetizar las desviaciones normalizadas de toda una tabla sumándolas, es preciso elevarlas al cuadrado para evitar que se anulen. De todo esto la fórmula del chi-cuadrado resulta en:

$$\chi^2 = \Sigma (f_o - f_e)^2 / f_e$$

Los estadísticos justamente aconsejan no calcular χ^2 si en la tabla hay al menos una frecuencia teórica igual a 5 o menor. Esto se debe a que en la fórmula cada sumando $(f_o - f_e)^2$ está normalizado dividiéndolo por f_e; si el divisor (f_e) es demasiado pequeño, el valor del sumando es demasiado alto, lo que inflará artificialmente el valor de la suma —es decir, del mismo χ^2.

Como ya se dijo, el chi-cuadrado no cuantifica la intensidad de la asociación entre dos variables categoriales. El lector puede comprobarlo calculando (con la ayuda de la computadora) su valor en una tabla, luego duplicando todas las cifras en la misma tabla y calculando nuevamente el coeficiente: su valor será exactamente el doble del valor precedente, mientras la intensidad de la asociación permanece igual. Por otro lado, es intuitivo que la atracción / repulsión que aparece en la tabla con las cifras duplicadas sea dos veces menos probablemente debida al azar porque está basada en el doble de casos.

El hecho de que el valor del chi-cuadrado dependa linealmente del número de casos demuestra su naturaleza de medida inferencial, es decir relativa a la inferencia de los resultados del análisis de una muestra a toda la población. Formalmente, para convertirlo en una medida relativa a la intensidad de la asociación basta con dividir su valor por N, el número de casos. Y en efecto, esto es lo

[4] Véase más adelante en este mismo apartado.

[5] Nótese que 'chi' es la transliteración inglesa de la letra griega χ. Las malas traducciones al español de los manuales norteamericanos de estadística impusieron el uso de 'chi' en lugar de 'ji', que sería la transliteración correcta en lengua española.

[6] Es decir, no existen atracciones ni repulsiones entre las categorías de una variable y las categorías de la otra: véase el apartado 15.2.

que hace la fórmula de φ^2 (Fi-cuadrado); también otros dos coeficientes, C de Pearson y V de Cramér, se basan en esta división, en la que introducen varias modificaciones.

Pero, como se decía, coeficientes que sinteticen la intensidad de una asociación entre variables categoriales no tienen sentido porque las relaciones se sitúan entre cada categoría de una variable y cada categoría de la otra, y no pueden ser sintetizadas de ninguna forma. Aun cuando sea más rápido y parezca científico mencionar números, ningún coeficiente puede reemplazar la paciente confrontación entre frecuencias empíricas y frecuencias teóricas en cada celda.

Además, esta confrontación se debe hacer teniendo en cuenta lo que pasa en las otras celdas de la misma fila y de la misma columna: sólo así se pueden interpretar con mayor confianza los resultados (es decir, las frecuencias empíricas) en las celdas con bajas frecuencias teóricas, que —como ya se señalaba— no bastarían para establecer una conclusión si son consideradas aisladamente.

15.4. La relación entre dos variables ordinales

Como se decía en el apartado 7.2, una variable ordinal tiene valores altos y valores bajos. Si se relaciona con otra variable ordinal,[7] y a valores altos de la primera tienden a corresponder valores altos de la segunda, mientras a valores bajos de la primera tienden a corresponder valores bajos de la segunda, se dice que entre las dos variables existe una COGRADUACIÓN (también denominada concordancia).

La forma más rápida de evaluar qué relaciones existen entre las dos variables ordinales es considerar las celdas ubicadas en o cerca de las diagonales[8] de la tabla bivariable. Si —como es adecuado hacer— ponemos los valores bajos de ambas variables en el rincón izquierdo superior y los valores altos de ambas variables en el rincón derecho inferior, y encontramos desviaciones positivas en las celdas en o cerca de la diagonal, se puede concluir que las dos variables están asociadas positivamente. La tabla 15.11 presenta un caso.

[7] Si una variable ordinal se relaciona con una variable categorial (o con una dicotomía) la herramienta para analizar la relación es la tabla bivariable, que se analiza con los mismos criterios introducidos en el apartado precedente, y disfrutando el hecho de que la reducida autonomía semántica de las variables ordinales (véase apartado 8.2) permite a veces identificar un rumbo más general en la relación.

Por el contrario, si una variable ordinal se relaciona con una variable cardinal, la técnica de análisis cambia completamente (véase el apartado siguiente).

[8] En rigor, sólo se puede hablar de una diagonal si las dos variables tienen el mismo número de categorías. Pero, sea cual fuere la forma de la tabla, siempre se puede trazar una diagonal ideal y ver cuáles celdas son vecinas.

Tabla 15.11: *Un caso de asociación entre dos variables ordinales*

TIPO DE ÁREA DE RESIDENCIA	TÍTULO DE ESTUDIO NINGUNO	PRIMARIA	SECUNDARIA	LICENCIATURA	DOCTORADO	TOTAL
de montaña	10	15	15	4	1	45
rural	14	25	36	14	3	92
semirrural	6	22	38	14	3	83
de cinturón	7	29	44	21	3	104
urbano	13	46	99	61	16	235
capital provincial	12	45	98	72	22	249
metrópolis	6	38	65	60	23	192
total	68	220	395	246	71	1.000

La tabla muestra la relación entre el título de estudio y el tipo de área donde habitan los encuestados. Las categorías de la segunda variable están ordenadas desde la más marginal (área de montaña) a la más central. Se pusieron en cursiva las celdas donde la frecuencia empírica es mayor que la teórica. Se puede comprobar que todas esas celdas están ubicadas a lo largo de la diagonal de la cograduación (concordancia), la que desciende de la celda superior izquierda (valores bajos / bajos) a la inferior derecha (valores altos / altos). Por eso se puede concluir que la centralidad del área de residencia y el título de estudio están asociados: al aumentar la primera tiende a aumentar el segundo.

Naturalmente, no siempre pasa que las celdas con frecuencias empíricas más altas que las teóricas se ubiquen a lo largo de la diagonal descendente, como en la tabla 15.11. Esas celdas pueden ubicarse a lo largo de la diagonal ascendente (la que va de la celda inferior izquierda a la superior derecha): en este caso las dos variables ordinales estarían contragraduadas (discordantes), es decir tendrían una relación global negativa entre sí. O bien ellas pueden estar dispersas sin ningún patrón en la tabla; en este caso no se puede hablar de una relación global entre las dos variables, sino de un complejo de relaciones de atracción y repulsión entre categorías específicas de una y otra variable.

Los coeficientes que cuantifican la intensidad de la asociación entre dos variables ordinales asumen que la relación sea global: de lo contrario, los coeficientes tendrán valores bajos y poco sentido. Todos esos coeficientes se basan en la diferencia entre el total de pares "concordantes" y el total de pares "discordantes". Un par de casos (A y B) se denomina concordante cuando el caso A tiene un puntaje más alto que el caso B en ambas variables; se denomina discordante cuando el caso A tiene un puntaje más alto que el caso B en una variable y más bajo en la otra. Una tercera posibilidad es que dos casos tengan el mismo valor en una o ambas variables: este par de casos se denomina "empatado".

El más conocido de los coeficientes que cuantifican la intensidad de la asociación entre variables ordinales es γ (gamma: una letra griega correspondiente a una g gutural). Como otros coeficientes ordinales, γ se basa en la diferencia entre el total de pares concordantes (P en las fórmulas) y el total de pares discordantes (Q en las fórmulas), no considerando los pares empatados. Su fórmula es:

$$\gamma = \frac{P - Q}{P + Q}$$

Al no considerar los pares empatados, γ tiende a sobrevalorar la intensidad de la asociación, positiva o negativa; esta sobrevaloración es más grave al crecer el número de categorías de una u otra variable (Galtung 1967: 223; Davis 1971: 75-76).

Para poner remedio a las limitaciones de γ, Kendall propuso (1948, cap. 4) otros dos coeficientes para variables ordinales. Uno de ellos (τ_b, tau$_b$) puede conseguir el máximo valor (+1 o −1) sólo si la tabla bivariable es cuadrada (tiene el mismo número de categorías en las dos variables). El coeficiente tiene en el denominador, en lugar de la suma de pares concordantes y discordantes (como γ), el total de los pares posibles [N * (N − 1) / 2]. Su fórmula es:

$$\tau_b = \frac{2\,(P - Q)}{N\,(N - 1)}$$

Otro coeficiente (τ_c, tau$_c$) puede alcanzar su máximo valor (+1 o −1) en cualquier tipo de tabla.[9] Su fórmula es:

$$\tau_c = \frac{2 * \min * (P - Q)}{N^2\,(\min - 1)}$$

donde min indica el número menor entre el total de las filas y el total de las columnas de la tabla. Pero si la tabla no es grande, la disparidad entre min (que está en el numerador) y el valor de min −1 (que está en el denominador) puede conllevar un valor global del coeficiente τ_c que sobreestime la intensidad de la asociación.

Todos estos coeficientes son bidireccionales. Tal como hizo para las dicotomías, Somers propuso un coeficiente unidireccional, que llamó D como otros coeficientes que también había propuesto. Su fórmula es:

$$D = \frac{P - Q}{P + Q + V_d}$$

donde V_d indica el total de los pares que tienen el mismo valor en la variable considerada dependiente.

La tabla 15.12 presenta en forma sintética las varias situaciones en que se debe usar uno u otro coeficiente.

[9] No confundir con t_C que se presentó en el apartado 15.2 y que se refiere a dicotomías. Dado que la fantasía de los estadísticos no es grande, hay una media docena de coeficientes llamados t, que sólo se distinguen entre sí por el subíndice. Además no hay acuerdo entre autores en torno de los subíndices: el mismo subíndice es usado para coeficientes diferentes, y el mismo coeficiente es identificado con subíndices diferentes por autores diferentes. La única manera de salir de este caos es mirar las fórmulas.

Tabla 15.12: *Criterios para seleccionar un coeficiente de asociación entre variables ordinales*

LA RELACIÓN ES CONSIDERADA:	LA TABLA TIENE FORMA:	EL NÚMERO MENOR DE CATEGORÍAS ES:	EL COEFICIENTE ADECUADO ES:
bidireccional	cuadrada	cualquiera	τ_b
bidireccional	rectangular	< 6	γ
bidireccional	rectangular	≥ 6	τ_c
unidireccional	cualquiera de las dos	cualquiera	D

Las variables ordinales que se acostumbra usar en las ciencias sociales raramente tienen más que una media docena de categorías. Pero en los laboratorios de psicometría, y también en algunas encuestas de *marketing*, se usa otro tipo de variable ordinal, que se produce pidiendo a los sujetos ordenar un conjunto de objetos (pueden ser productos, actores, personajes políticos, etcétera) según un criterio escalar (de mayor a menor utilidad, o simpatía, etcétera). Para cuantificar el grado de concordancia entre los ordenamientos de los mismos objetos propuestos por sujetos diferentes, o por el mismo sujeto con criterios diferentes (por ejemplo, dos ordenamientos de los mismos políticos en orden de simpatía y de capacidad) el psicólogo inglés Charles Spearman propuso (1904b) el coeficiente ρ (rho, una letra griega correspondiente a una r aspirada). Su fórmula es:

$$\rho = 1 - \frac{6 \sum d^2}{N (N^2 - 1)}$$

donde d es la diferencia en las posiciones del mismo objeto en los dos ordenamientos y N es el número de objetos.

La fórmula es concebida de manera que r cobra su máximo valor positivo (+1) cuando hay coincidencia perfecta entre los ordenamientos, y su máximo valor negativo (–1) cuando un ordenamiento es exactamente inverso al otro.

Rho ocupa una posición intermedia entre los coeficientes ordinales y los cardinales (de los que se hablará en el apartado siguiente). No es un coeficiente ordinal porque los coeficientes ordinales no tratan en ninguna forma de cuantificar la distancia entre posiciones, limitándose a establecer si un valor es mayor, igual o menor que otro. Y no es un verdadero coeficiente cardinal porque la distancia entre las posiciones es *supuesta* igual pero no es medida de ninguna forma. Por eso ρ, al que se denomina coeficiente de rangos, sólo se puede aplicar a la concordancia entre ordenamientos: cualquier otro uso *sería incorrecto*.

15.5. La relación entre una variable cardinal y una categorial u ordinal

La forma más simple de relacionar una variable cardinal con una categorial es presentar algunos valores característicos (como el promedio, la desviación típica, etcétera) de la cardinal para cada categoría de la categorial. La tabla 15.13 presenta esos valores relativos al ingreso mensual (en dólares) de ciudadanos de cuatro países.

Tabla 15.13: *Promedio, desviación típica y coeficiente de variación del ingreso mensual en varios países*

PAÍS	PROMEDIO (U$S)	DESVIACIÓN TÍPICA (U$S)	COEFICIENTE DE VARIACIÓN
Finlandia	2.538	824	0,325
Suecia	2.404	607	0,252
México	827	548	0,663
Uruguay	811	463	0,571

A pesar de su simpleza, la tabla permite sacar conclusiones interesantes en torno del nivel de ingreso medio cotejado con su dispersión. Se comprueba por ejemplo que los ciudadanos de Suecia y de Finlandia tienen un nivel similar de ingreso medio, pero entre los suecos hay una mayor homogeneidad en el ingreso, probablemente debido al hecho de que Finlandia tiene una población más numerosa en áreas rurales pobres. Lo mismo ocurre para mexicanos y uruguayos: ingreso promedio similar, pero mayor homogeneidad entre los últimos.

La (probablemente alta) diferencia de ingreso entre un sueco cualquiera y un mexicano cualquiera depende en alguna medida de la variable categorial, es decir, del país de residencia (quien vive en Suecia tiene a priori cierta ventaja), mientras que la diferencia de ingreso entre un mexicano cualquiera y un uruguayo cualquiera no depende en medida apreciable del país de residencia, sino de otros factores (profesión, educación, etcétera).

Una figura como la 15.3 (denominada *BOX PLOT*, diagrama de cajones) es el equivalente gráfico de una tabla como la 15.13 respecto de informaciones acerca del promedio y dispersión de una variable cardinal entre las categorías de una variable categorial. En particular, en la figura 15.3 se representan los ingresos mensuales (en dólares) declarados por empleados encuestados en cinco países de Sudamérica: Argentina, México, Uruguay, Brasil, Colombia. Los *boxes* unen para cada país el primer y el tercer cuartil.[10] Al interior de cada *box*, una línea espesa indica la posición de la media (o de la mediana:[11] eso depende de la elección de quien diseña el *box plot*). En muchos *box plots* se ponen también los *whiskers* (favoritos), líneas que unen el primer cuartil con el primer decil[12] o el tercer cuartil con el último decil. Si el número de casos no es alto (como en el caso de los encuestados uruguayos en la figura 15.3) puede pasar que un decil tenga el mismo valor del cuartil relativo, y entonces no se traza. Los circulitos que se ven en la figura 15.3 son letras o, que representan *outliers*,[13] e indican casos con valores muy diferentes de los otros.

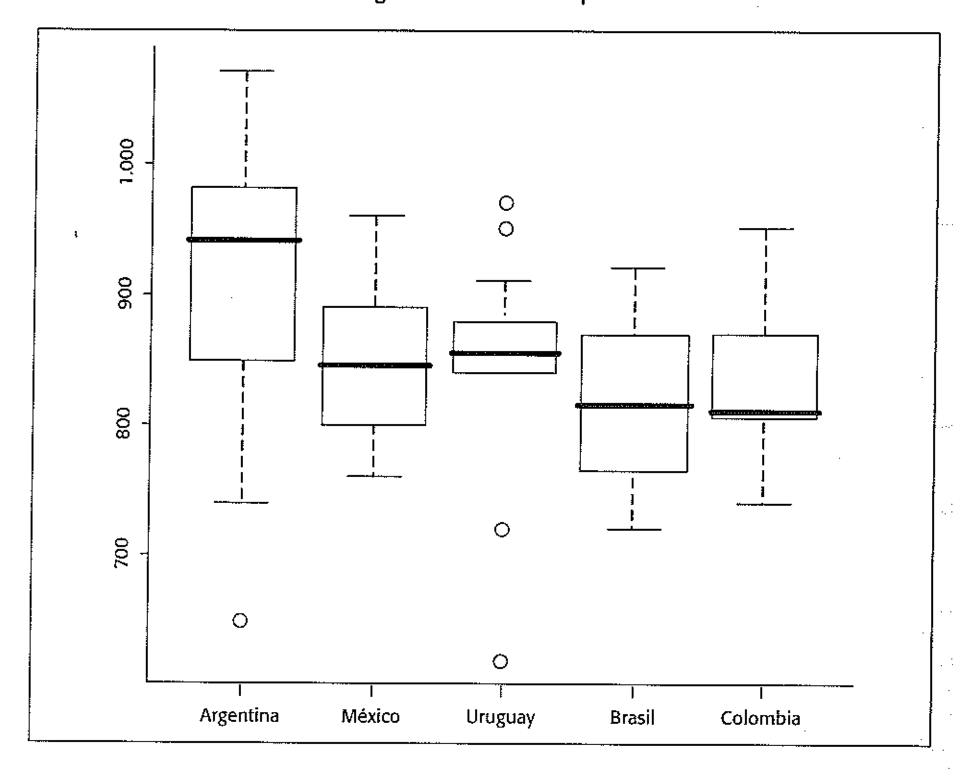
[10] En una distribución el primer cuartil es el valor que divide el cuarto de los casos con valores más bajos que todos los otros; el tercer cuartil es el valor que divide el cuarto de los casos con valores más altos que todos los otros.

[11] La mediana es la cifra que divide la distribución en dos partes iguales: una mitad tiene los valores más bajos, otra tiene los valores más altos (véase apartado 8.2).

[12] En una distribución el primer decil es el valor que divide el diezmo de los casos con valores más bajos que todos los otros; el último decil es el valor que divide el diezmo de los casos con valores más altos que todos los otros.

[13] *Outlier* es un término inglés que indica algo que yace (*lies*) afuera (*out*) del grupo.

Figura 15.3: *Un* box plot



En la figura se comprueba que el promedio de los ingresos mensuales declarados por los empleados argentinos es más alto que el de los otros cuatro países; por lo que concierne a los empleados argentinos respecto de otros, la variable categorial 'país de residencia' tiene cierta influencia. Por otro lado, los ingresos mensuales declarados por los empleados de los otros cuatro países tienen un promedio equivalente: los uruguayos se distinguen no tanto por tener un promedio más alto, sino por declarar ingresos más homogéneos que otros.[14]

Imaginemos ahora reiterar el mismo análisis para los encuestados de un solo país, considerando en este caso como categorías a las provincias de residencia. Entonces, se divide la varianza total de los ingresos[15] en dos partes: una parte se puede atribuir a la pertenencia de cada encuestado a una u otra provincia, y la parte restante se debe atribuir a otros factores. La segunda parte se denomina varianza interior (a cada categoría); la primera se denomina varianza exterior; por eso la técnica se denomina ANÁLISIS DE LA VARIANZA. El coeficiente que cuantifica la influencia de la variable categorial en la cardinal está basado en la proporción de la varianza exterior sobre la varianza total. Esto fue propuesto

[14] Con la excepción de cuatro *outliers*; el más bajo de ellos podría ser una respuesta maliciosa o un error en la codificación.

[15] Sobre el concepto de varianza, véase el apartado 8.3.

por Karl Pearson (1913) y su símbolo es η^2 (eta cuadrado: una letra griega correspondiente a una e larga). Su fórmula es:

$$\eta^2 = \sum_{i}^{N} \sum_{j}^{k} x^2_{.j} \Big/ \sum_{i}^{N} \sum_{j}^{k} x^2_{ij}$$

Como se vio en el apartado 8.3, x_i indica la desviación de cualquier dato en la variable cardinal X del promedio de la misma variable; $x_{.j}$ indica que las desviaciones conciernen no a casos específicos (porque el símbolo de los casos específicos, el subíndice *i*, está neutralizado y sustituido por un punto), sino a la media de la variable cardinal en las distintas categorías de la categorial. En el denominador el signo x_{ij} indica que en este caso se consideran las desviaciones de cada caso en cada categoría del promedio general.

Hay dos símbolos de sumatoria (Σ: véase el apartado 8.3) para destacar que los promedios de las categorías no deben sumarse una vez por cada categoría, sino que cada promedio debe sumarse tantas veces cuantos sean los casos pertinentes en su categoría.

El valor de η^2 varía entre 0 (el promedio de la variable cardinal es exactamente el mismo en todas las categorías: eso demuestra que la variable categorial no tiene ninguna influencia sobre la cardinal) y 1 (todos los casos pertinentes a una categoría tienen el mismo valor en la variable cardinal: no hay varianza interior a las categorías, y toda la varianza depende de la pertinencia a una u otra categoría de la variable categorial). Entre esos dos extremos se ubican los valores de η^2: habitualmente, en las variables que se utilizan en las ciencias sociales, ellos no superan .2, y un η^2 igual a .1 ya se puede considerar la señal de una influencia considerable de la variable categorial en la cardinal.

Cabe destacar el hecho de que cuanto mayor es el número de las categorías de una variable categorial, mayor es la probabilidad de que una parte de la varianza total de la cardinal se convierta en varianza exterior (entre una categoría y la otra). Por ese motivo, si dos variables categoriales tienen un número sensiblemente diferente de categorías no es preciso confrontar los η^2 para ver cuál de las dos influye más a una variable cardinal: la que tiene más categorías goza de una ventaja *a priori*, independiente de su influencia real. Por ejemplo, si en la Argentina la asociación entre género e ingreso tiene el mismo η^2 que la asociación entre ingreso y provincia de residencia, la conclusión es que el género tiene una influencia mayor que la provincia de residencia.

Si una variable cardinal (por ejemplo, el ingreso) se relaciona con una variable ordinal (por ejemplo, el título de estudio), se puede esperar que se halle una asociación global, en el sentido de que el promedio de ingreso en cada categoría de la variable ordinal crezca a medida que los títulos sean superiores (si estos promedios disminuyeran regularmente al ser mayor el título, la asociación sería igualmente global, pero negativa en vez de positiva).

Si los promedios de la variable cardinal en las categorías de la ordinal crecen o disminuyen regularmente, se dice que hay una relación monotónica en-

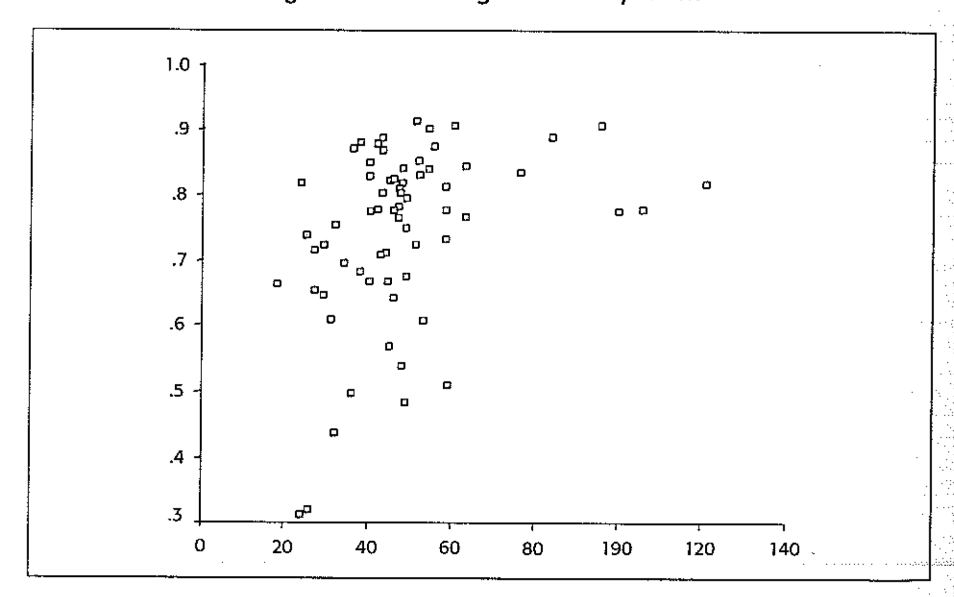
tre las dos variables.[16] Naturalmente, una figura como el *box plot* ayuda a comprobar si la relación entre las dos variables es global (en tal caso, los promedios de la cardinal en las categorías crecen o disminuyen regularmente) o si algunas categorías tienen asociaciones particulares con la cardinal, en la medida que sus promedios son considerablemente más altos o más bajos de lo que deberían ser siguiendo el orden.

Una medida de cuánto se aleja de ser global la asociación entre las dos variables se puede lograr restando de η^2 el coeficiente equivalente para dos variables cardinales (r^2), que asume una relación global —como se verá en el próximo apartado. Si la diferencia es grande, eso significa que los promedios de la variable cardinal en una o más categorías se apartan mucho del rumbo de las otras porque esas categorías tienen relaciones particulares (llamadas "interacciones") con la variable cardinal.

15.6. La relación entre dos variables cardinales

La relación entre dos variables cardinales se representa poniendo una en la abscisa y otra en la ordenada de un diagrama de dispersión, como el que se ve en la figura 15.4. Por convención se pone en la abscisa y se indica con la letra X la variable que se considera independiente en el modelo (véase el apartado 15.1) y en la ordenada la que se considera dependiente, indicándola con la letra Y. Cada punto en el diagrama representa un caso, o más exactamente el cruce de sus valores en la variable que está en la abscisa y en la variable que está en la ordenada.

Figura 15.4: *Un diagrama de dispersión*



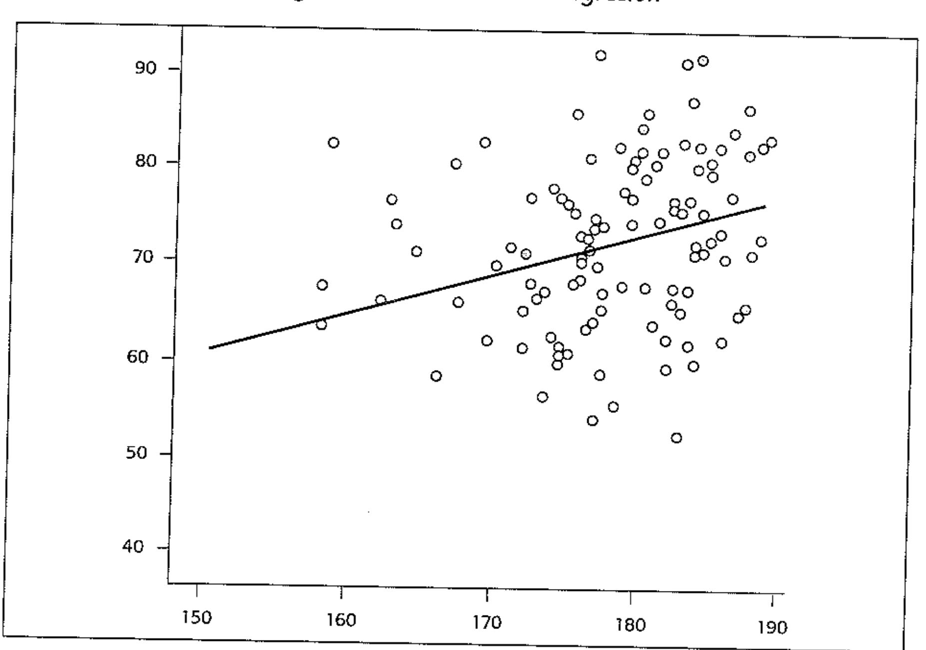
[16] No se puede hablar de una relación lineal porque una de las variables es ordinal, lo que conlleva que la distancia entre sus categorías es arbitraria: sólo el orden de las categorías no lo es.

En la figura 15.4 se presenta la relación entre el número de municipios de una provincia y la proporción de los que están conectados por el ferrocarril. La relación no es lineal: lo sería si al crecer el número de municipios creciera regularmente la proporción de los que están conectados por el ferrocarril. Las provincias de montaña tienen pocos municipios y muchos no están conectados, mientras que una proporción bastante alta de municipios en las provincias de llanura está conectada; pero algunas provincias de llanura son pequeñas y tienen pocos municipios, similar a las de montaña. La forma no lineal del patrón que aparece en el diagrama atestigua la naturaleza espuria[17] de la asociación entre el número de municipios de una provincia y la proporción de esos municipios que son conectados por el ferrocarril: lo que influye en la segunda variable no es el número de municipios en sí sino el ser la provincia de montaña o de llanura.

Si la posición de los puntos en el diagrama presenta una relación bastante lineal entre las dos variables (en el sentido de que los valores de la variable en la ordenada tienden a crecer, o decrecer, regularmente al crecer los valores de la variable en la abscisa) se puede cuantificar con varios coeficientes la influencia que la variable que se considera independiente ejerce en la que se considera dependiente.

El primero de ellos mide la pendiente de la recta que interpola los puntos del diagrama, como la que se ve en la figura 15.5, que presenta la relación entre la estatura y el peso de una muestra de encuestados. Naturalmente, el peso de un individuo no está influido sólo por su altura, sino por muchas otras propiedades que no aparecen en el modelo. Con todo, el diagrama atestigua cierta influencia de la estatura en el peso en el sentido de que, al crecer una, tiende a crecer el otro.

Figura 15.5: *Una recta de regresión*



[17] Sobre relaciones espurias, véase el capítulo siguiente.

La recta que interpola los puntos del diagrama en la figura 15.5 se denomina RECTA DE REGRESIÓN.[18] Alineados a lo largo de esta recta se encuentran los valores de la variable dependiente ($\hat{Y}_i$) que corresponderían a cada valor de la independiente (X_i) según la relación global entre X e Y. Ya que, como se dijo, la relación entre X e Y no es determinista, y los valores Y_i también son influidos por muchas otras propiedades, casi siempre hay un desfasaje entre los valores efectivos (Y_i) y los que la técnica prevé ($\hat{Y}$) sólo sobre la base de la asociación entre X e Y.

La diferencia entre Y_i e $\hat{Y}$ se denomina RESIDUO (e_i): por lo tanto,

$$Y_i = \hat{Y} + e_i$$

Figura 15.6: *Recta de regresión y residuos*

Como muestra la figura 15.6, en la que algunos residuos están evidenciados, los residuos son segmentos de la ordenada: cuando el valor efectivo Y_i es más alto del valor previsto ($\hat{Y}$), el residuo es positivo; a la inversa, es negativo. Como pasa con todas las desviaciones de un promedio, la suma algebraica de los residuos positivos y negativos es 0. El lector no debe sorprenderse por esta afirmación: la recta de regresión puede considerarse una media móvil; cada punto en esta recta puede concebirse como el promedio de todos los valores que podría asumir la Y en correspondencia con un valor dado de la X si la relación global entre X e Y fuese la trazada por aquella recta de regresión.

Veamos ahora cómo se traza esta recta de regresión. Para ello, necesitamos dos coeficientes: uno es *a*, denominado intersección, que es el valor previsto de Y donde la recta de regresión intercepta la ordenada, cuando la variable independiente X tiene valor 0. Otro coeficiente es *b* (denominado "coeficiente de regresión"), que mide la pendiente de la recta, es decir el incremento del valor previsto de Y para cada incremento unitario de los valores X_i.

[18] Una etiqueta que le dio el gran estadístico inglés Francis Galton por sugerencia de su colega Francis Ysidro Edgeworth (Piovani 2006).

La ecuación de la recta es por lo tanto:

$$\hat{Y}_i = a + bX_i$$

donde $\hat{Y}_i$ significa —como ya se dijo— valor previsto de Y_i en base a la relación global de Y con X.

En la fórmula, *a* y *b* son incógnitas. Para establecer su valor con respecto a los datos que estamos examinando, se parte de lo que ya se dijo, es decir que los residuos pueden considerarse desviaciones de un promedio y que por lo tanto su suma algebraica es 0. Esto se expresa con la equivalencia $\sum (Y_i - \hat{Y}_i) = 0$.

Pero la fórmula anterior establece que $\hat{Y}_i = a + bX_i$. Por lo tanto se puede expresar $\sum (Y_i - a - bX_i) = 0$. Dado que *a* y *b* son constantes, distribuyendo la sumatoria esta equivalencia se puede expresar $\sum Y_i - Na - b\sum X_i) = 0$, donde N es el número de casos.

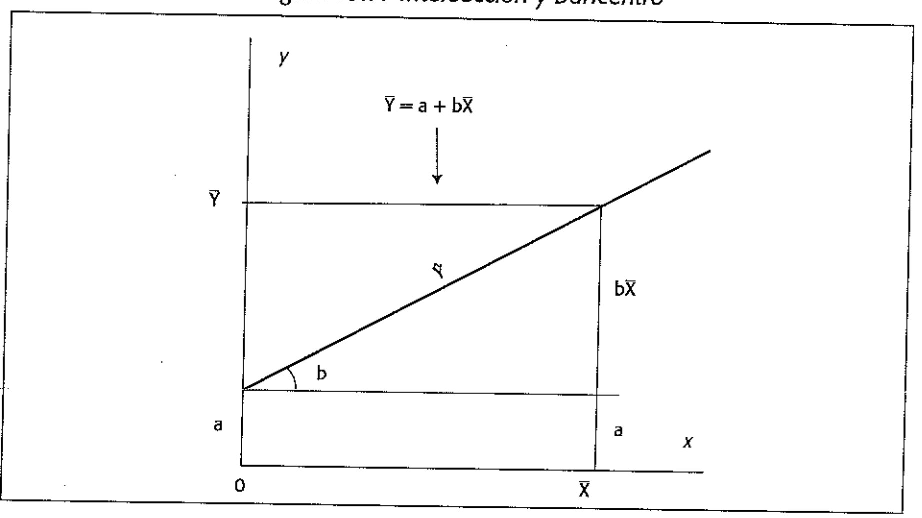
Pasando Na al otro lado del igual con un cambio de signo, obtenemos:

$$Na = \sum Y_i - b\sum X_i$$

y dividiendo ambos miembros por N obtenemos: $a = \hat{Y} - b\hat{X}$.

Este resultado, obtenido tras manipulaciones matemáticas de la fórmula de partida, nos dice que el valor de la INTERSECCIÓN es igual a la diferencia entre el promedio de los valores de la variable dependiente y el promedio de los valores de la variable independiente multiplicados por el coeficiente de regresión. Esto conlleva que la recta de regresión tiene que pasar —como lo muestra la figura 15.7— por el punto del diagrama donde se cruzan la ordenada de $\hat{X}$ y la abscisa de $\hat{Y}$. Este punto se denomina BARICENTRO, porque corresponde al punto de equilibrio del sistema de los puntos.

Figura 15.7: *Intersección y baricentro*

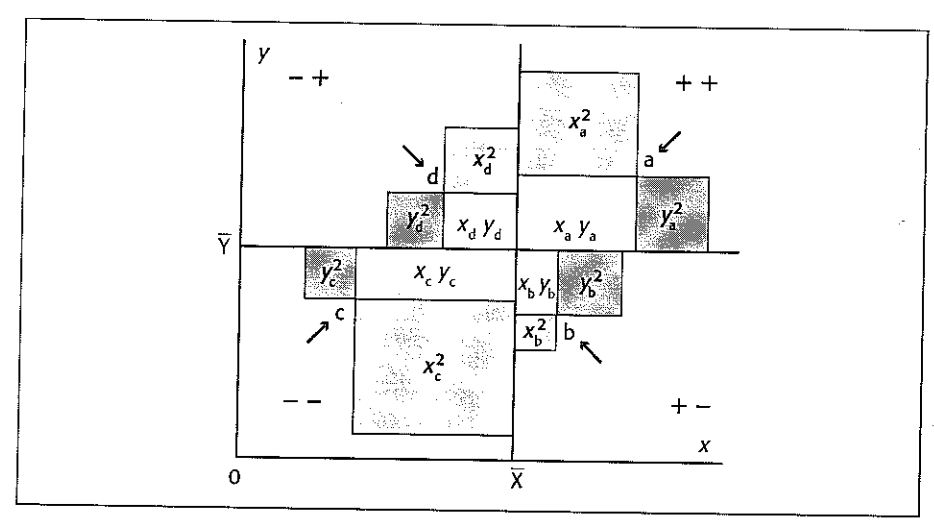


Pero, como lo establece la geometría, por un punto pasan infinitas rectas. Para saber cuál es la recta de regresión tenemos que conocer el valor de *a* (porque por dos puntos pasa una sola recta), y para calcular éste con la fórmula: $a = \bar{Y} - b\bar{X}$ tenemos que conocer $b\bar{X}$. Para determinar *b*, y por consiguiente *a*, se necesita introducir otra condición: que la suma de los residuos cuadrados sea la menor posible. Con este criterio[19] se fuerza la recta a pasar lo más cerca posible a la mayor cantidad de puntos en el diagrama, ya que la elevación al cuadrado acrecienta el peso de los residuos más grandes. Aplicando la fórmula del análisis matemático para la minimización, se halla que el valor de *b* que logra este resultado es

$$\text{b (coeficiente de regresión)} = \frac{\sum x_i y_i}{\sum x_i^2}$$

El coeficiente de regresión de Y en X es por lo tanto la codevianza de X e Y (en cuánto las dos variables varían juntas) dividido por la devianza de X (en cuánto la independiente varía sola). La división por la devianza tiene el efecto de normalizar *b* con respecto al número de casos, en cuanto este número influye en ambas sumas (tanto la del numerador como la del denominador). La fórmula tiene una interpretación geométrica, que se presenta en la figura 15.8.

Figura 15.8: *Interpretación geométrica de los coeficientes de regresión y correlación*



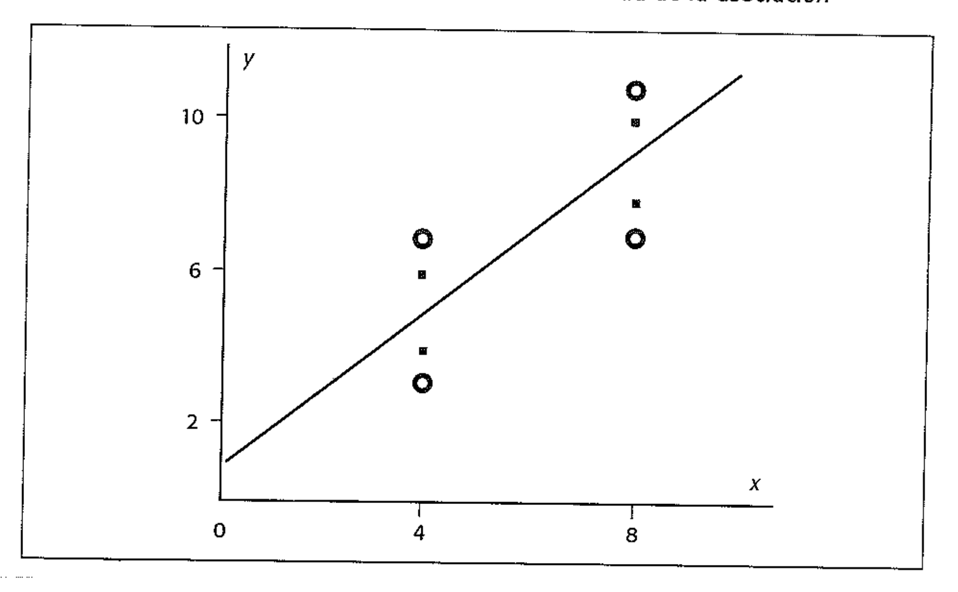
En el diagrama de la figura 15.8 sólo se ubicaron cuatro puntos (uno por cada cuadrante) señalados por flechas. Para cada punto se diseñaron dos desviaciones: un segmento horizontal representa su desviación del promedio $\bar{X}$, y un segmento vertical representa su desviación del promedio $\bar{Y}$. Estos dos segmentos forman, con la abscisa y la ordenada de los promedios, un rectángulo (blanco en la figura 15.8) que es la representación gráfica del sumando $x_i y_i$ en el numerador de la fórmula de arriba. Además, sobre cada desviación en la abscisa se construyó un rectángulo con fondo grisado claro, que es la representación gráfica del sumando x_i^2 en el denominador de la fórmula enunciada.[20] El coeficiente de regresión es nada más que el cociente entre la suma de los rectángulos blancos dividida por la suma de los cuadrados con grisado claro (naturalmente en los diagramas con datos reales hay mucho más que cuatro puntos).

[19] Introducido hace dos siglos por los matemáticos Gauss y Legendre.

Se dijo que *b* está normalizado con respecto del número de casos; pero no lo está con respecto a la unidad de medida de las dos variables, en tanto que en el denominador de la fórmula sólo está X. Si la variable en la abscisa fuera alguna longitud en kilómetros, convirtiendo la medida en metros automáticamente se dividiría por 1.000 el coeficiente de regresión de cualquier otra variable sobre la longitud. El coeficiente de regresión sólo sirve para prever el valor de cada caso en la variable Y, si se conoce el valor de X en ese caso y teniendo en cuenta la intersección y las específicas unidades de medida de X e Y. Pero no sabemos nada acerca de cuánto esta revisión sea confiable, en tanto no sabemos nada acerca de la intensidad de la asociación entre X e Y.

Se intuye que cuanto más cercanos son los puntos a la recta de regresión, tanto más confiable es la previsión; pero *b* mide la pendiente de la recta, no la distancia de los puntos a ella. La figura 15.9 y la tabla 15.14 proporcionan una demostración de esto, presentando la relación con la misma X (en la abscisa) de dos variables dependientes, cuyos valores son representados en la ordenada por cuatro cuadraditos negros (variable Y) y por cuatro circulitos (variable Y').

Figura 15.9: *Insensibilidad de* b *a la intensidad de la asociación*



[20] Los cuadrados construidos sobre las desviaciones en ordenada (grisado oscuro) no se consideran para el coeficiente de regresión, sino para el de correlación.

La figura hace patente que la asociación de Y con X es más estrecha, ya que los cuadraditos negros están más cercanos a la recta de regresión, lo que hace más confiable la previsión y más pequeños los residuos. Con todo, la pendiente de la recta de regresión es la misma para las dos variables (Y e Y').

Esto se puede comprobar con la tabla 15.14. En la primera columna se reportan los valores de los cuadraditos negros en la abscisa; en la segunda, los cuadrados de sus desviaciones (símbolos x_i^2) del promedio (que es 6). En la tercera y la cuarta columna, se repite lo mismo para la ordenada (cuyo promedio es 7). En la quinta columna se reportan los productos de las desviaciones en la abscisa y en la ordenada.

Tabla 15.14: *Datos relativos a la figura 15.9*

← (CUADRADITOS) →						← (CIRCULITOS) →		
X_i	x_i^2	Y_i	y_i^2	x_iy_i	e_i	Y_i'	$y_i'^2$	x_iy_i'
4	4	4	9	6	−1	3	16	8
4	4	6	1	2	+1	7	0	0
8	4	8	1	2	+1	7	0	0
8	4	10	9	6	−1	11	16	8
24	16	28	20	16	0	28	32	16
prom.	↑	prom.	↑	↑		prom.	↑	↑
6	DEV_X	7	DEV_Y	$CODEV_{X,Y}$		7	$DEV_{Y'}$	$CODEV_{X,Y'}$

El total de la segunda columna es la devianza de X; el total de la quinta columna es la codevianza de X e Y. Dividiendo la codevianza (16) por la devianza de X (16) se obtiene 1. Por consiguiente la intersección, que es igual a $\bar{Y} - b\,\bar{X}$, es (7 − 6 =) 1. Dado que la pendiente y la intersección son igual a 1, sumando 1 a cada X_i podemos calcular los $\bar{Y}$ correspondientes; restando esos valores de los Y_i efectivos (tercera columna) se obtienen los residuos, que se presentan en la sexta columna.
Dado que los circulitos y los cuadraditos tienen iguales valores en la abscisa, las dos primeras columnas de la tabla 15.4 también sirven para los primeros. Las ordenadas de los circulitos (Y_i') se presentan en la séptima columna; los cuadrados de sus desvíos del promedio de los Y_i' (que es 7) se presentan en la octava columna, y la novena presenta los productos de los desvíos de cada circulito en la abscisa y la ordenada. Se puede comprobar que la codevianza de X e Y' es igual a la codevianza de X e Y. A su vez, la devianza de las X es la misma porque las abscisas son iguales; luego también *b* es igual, como lo muestra la figura 15.14.

La forma de obtener un coeficiente que no dependa de las unidades de medida y que mida la intensidad de la asociación es simple: basta con hacer totalmente especular la fórmula, añadiendo lo que faltaba al denominador, es decir la devianza de la variable dependiente. Así se obtiene el COEFICIENTE DE CORRELACIÓN, también denominado *r* de Pearson (en honor al estadístico que lo perfeccionó en 1896); su fórmula es

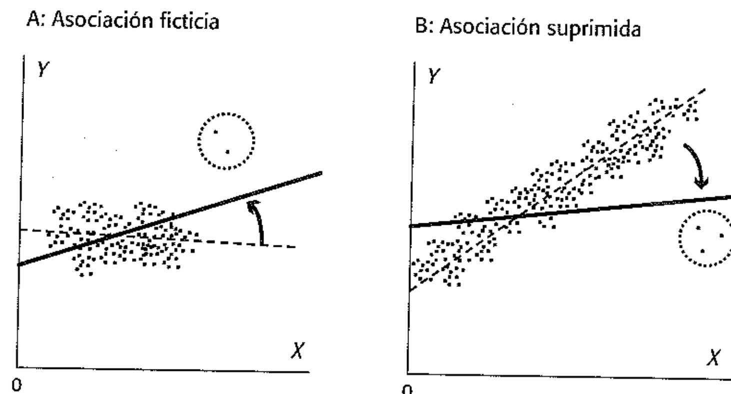
$$r = \frac{\sum x_i y_i}{\sqrt{\sum x_i^2 \sum y_i^2}}$$

El numerador de *r* es igual al de b; la diferencia está en el denominador, que usa la media geométrica de las dos devianzas en lugar de solamente la devianza de la independiente. Esta forma de normalización convierte a *r* en un coeficiente estandarizado, que varía entre un máximo de +1 (las dos variables coinciden en su dirección) y un mínimo de −1 (las dos variables siguen direcciones inversas). Si las dos variables son independientes, o entre ellas no hay alguna forma de relación lineal,[21] *r* cobra valor 0.

Siendo estandarizado, *r* no sirve (como *b*) para prever los valores a partir de los X_i; pero sí sirve para estimar la precisión de esa previsión. Además, sirve para comparar la intensidad de la asociación entre dos variables cualesquiera, y naturalmente de la asociación de una misma pareja de variables en poblaciones o muestras diferentes.

Los coeficientes de regresión y de correlación asumen que la relación entre X e Y sea global y lineal. Si una relación no lo es, los valores de esos coeficientes pueden ser gravemente sesgados. Las figuras 15.10A y 15.10B muestran dos situaciones en las que una relación no se puede definir como global, en el sentido de que la posición de algunos puntos en el diagrama muestra que en los casos correspondientes la relación entre X e Y es muy diferente de lo que es en otros casos.

Figura 15.10: *Efectos de* outliers



La figura 15.10A muestra una asociación ligeramente negativa, casi inexistente entre X e Y. Pero hay dos puntos muy alejados de los otros en el diagrama: ellos se encuentran fuera del grupo, y por eso se denominan *outliers.* Sus grandes desviaciones del promedio producen grandes sumandos que influyen en la fórmula de *b.* La recta de regresión, que sería ligeramente negativa, es atraída por esos *outliers* (como muestran las flechas en la figura), y

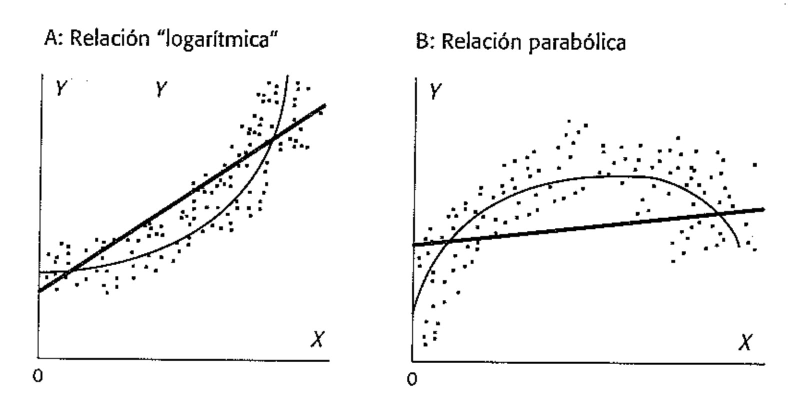
[21] Más adelante se aclarará el significado de esta frase.

el coeficiente *b* que mide su pendiente daría una impresión completamente sesgada de la naturaleza de la relación a quien no mire el diagrama.

Una distorsión opuesta muestra la figura 15.10B: la asociación es fuerte y positiva pero tres *outliers* atraen la recta de regresión. A quien no mirase el diagrama, sino sólo al coeficiente *b*, la asociación le parecería mucho menos intensa de lo que es efectivamente.

Por otro lado, la asociación puede ser global pero no lineal. La figura 15.11 presenta dos casos en que la asociación es más intensa y tiene una naturaleza diferente de lo que aparece con la recta de regresión y los coeficientes relativos. No sólo la naturaleza curvilínea de la asociación se pierde completamente (lo que es inevitable al interpolar una recta), sino también muchos residuos de la recta resultan más altos de lo que serían respecto de la curva. Luego, el valor de *r* resulta más bajo, detectando una asociación menos intensa de lo que es en realidad.

Figura 15.11: *Relaciones no lineales*



Naturalmente, los matemáticos han elaborado desde hace siglos las herramientas para interpolar curvas en lugar de rectas. Por ejemplo, la curva diseñada en la figura 15.11A tendría una ecuación de la forma $\hat{Y} = a + bX_i + cX_i^2$, y la curva diseñada en la figura 15.11B tendría una ecuación de la forma $\hat{Y} = a + bX_i - cX_i^2$. Estas fórmulas, y otras más complejas, se usan regularmente en las ciencias naturales y en la ciencia económica, en las que se puede razonablemente imaginar que algunos modelos bivariables estén bastante especificados (es decir, no estén demasiado alejados de las situaciones reales). Pero en las ciencias sociales los modelos bivariables sólo son la puerta de acceso (como se decía en el apartado 15.1) a modelos más complejos denominados "multivariables", que se realizan manipulando los coeficientes *b* o *r* relativos a las relaciones entre todos los pares de variables incluidas en cada modelo multivariable. Si estas relaciones bivariables son expresadas por más de un coeficiente, los modelos multivariables se convierten en complejos sistemas de ecuaciones denominadas "no lineales", que implican competencias matemáticas que no tiene la mayoría de los investigadores sociales.

Por ese motivo, en sociología y ciencia política se tiende a hipotetizar relaciones lineales entre propiedades. Si la inspección visual del gráfico muestra que las relaciones son patentemente curvilíneas (véanse las figuras 15.4 y 15.11A), la solución que se acostumbra adoptar es transformar los valores de una de las variables del modelo en sus logaritmos. Comprimiendo los valores en la abscisa o en la ordenada, esta transformación tiende a convertir en lineales relaciones que antes no lo eran. Naturalmente, una solución de ese tipo es atinada cuando la relación es "logarítmica" (como en la figura 15.11A) y no cuando es parabólica (como en la figura 15.11B).

Las relaciones parabólicas presentan problemas que no son solucionables a nivel elemental. Para poder aplicar modelos multivariables que incluyan una pareja de variables con una relación parabólica entre sí, algunos dividen los casos que están analizando en dos grupos: en uno ponen los casos en los que la asociación entre X e Y es creciente; en otro, los casos en los que es decreciente.

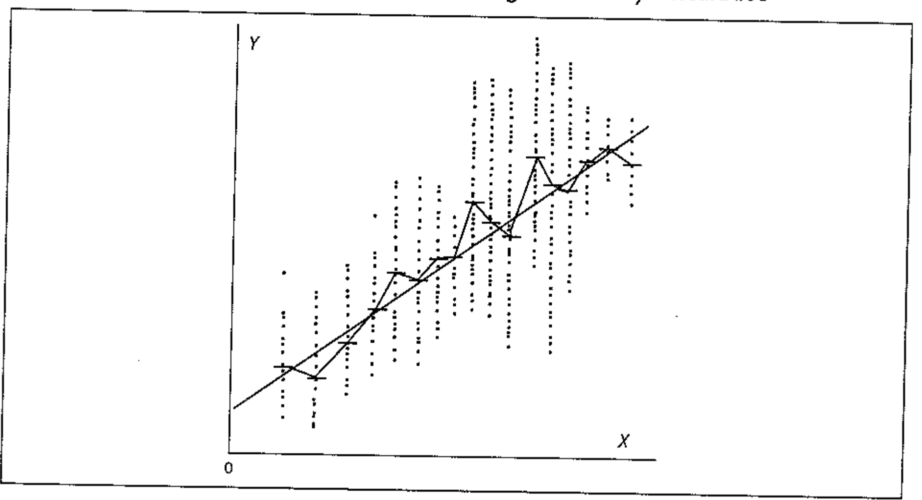
Hay una forma de cuantificar la medida en que la asociación entre dos variables cardinales o cuasi cardinales dista de ser lineal. Esa forma requiere del cuadrado de r (r^2), que se denomina COEFICIENTE DE DETERMINACIÓN, porque trata de medir cuánto los valores de la variable independiente determinan (linealmente) los valores de la dependiente. Este nivel de determinación lineal será tanto mayor cuanto más cercanos a la recta de regresión sean los puntos del diagrama.

El coeficiente r^2 es igual al cociente entre la devianza de Y reproducida[22] por su regresión lineal en X y la devianza total de Y. En fórmula:

$$r^2 = \frac{\Sigma(\hat{Y}_i - \bar{Y})}{\Sigma(Y_i - \bar{Y})}$$

El coeficiente r^2 es el equivalente de η^2 (véase apartado 15.5), que es adecuado cuando una de las dos variables es categorial u ordinal.

Figura 15.12: *Promedios de categoría libres y vinculados*



[22] En lugar de devianza (o varianza) reproducida se acostumbra hablar de devianza (o varianza) "explicada". Pero el verbo 'explicar' y sus derivados no parecen atinados para describir los resul-

Reflexionando sobre la figura 15.12 se entiende la naturaleza de la relación entre r^2 y η^2. En el diagrama se muestran con guiones los promedios de los Y_i correspondientes a cada X_i (cada valor de la X). En cambio, los valores de Y previstos por la recta de regresión de Y en X tienen que ser alineados a lo largo de esta recta. Ese vínculo se justifica a la luz del significado global que se acostumbra dar a la asociación entre dos variables cardinales, ya que en una variable cardinal el centro semántico es toda la variable. Por un motivo especular, el vínculo no es adecuado cuando la variable no es cardinal, y su centro semántico se ubica en cada categoría.

Para cuantificar la medida en que la asociación dista de ser lineal basta con tratar la variable independiente ya sea como cardinal, calculando r^2 (el coeficiente de determinación de los valores Y_i a partir de los valores X_i) ya sea como categorial, calculando η^2. A menos que todos los promedios representados por guiones estén alineados a lo largo de la recta (lo que no pasa en la figura 15.12), por necesidad matemática el valor de η^2 tiene que ser mayor del valor de r^2 porque, al no estar vinculados a la linealidad, los promedios de cada categoría se adaptan mejor a la distribución real de los puntos en el diagrama.

Cuanto más grande es la diferencia entre r^2 y η^2, más alejada de la linealidad está la asociación. Pero, a menos que la diferencia sea importante y sus motivos se puedan hipotetizar, cabe igualmente tratar a X como cardinal, conforme a su verdadera naturaleza.

Antes de terminar este largo capítulo, cabe mostrar cómo el análisis de los propios residuos que estadísticos y econometristas descartan como errores ofrece —en algunas situaciones— sugerencias para enriquecer el modelo, pasando de uno bivariable a uno multivariable a través de la incorporación de una o más variables que influyen en la dependiente.

Una condición esencial para lograr este resultado es que el investigador o los miembros de su equipo conozcan muy bien los casos y sean capaces de identificar cada uno de ellos en el diagrama. Lo que conlleva que esta ventaja del análisis de residuos se limita a investigaciones sobre variables cardinales relativas a agregados territoriales (las provincias de un Estado, los municipios de una provincia).[23] En este tipo de análisis pasa a menudo que en la misma área del diagrama (con abscisas y ordenadas similares y residuos todos positivos o todos negativos) se agrupen puntos correspondientes a provincias o municipios socioeconómicamente homogéneos. Al conocer las características de esas provincias o municipios no es difícil darse cuenta de cuál o cuáles son las propiedades que, influyendo en una dirección u otra en la variable dependiente, causan la presencia de residuos positivos o negativos en la regresión original.

Una vez identificadas, esta o estas variables pueden ser incluidas en modelos sucesivos, que así pasan de ser bivariados a ser multivariados.

---

tados de una elaboración matemática. Ésta puede sí cuantificar el nivel de asociación entre dos variables, pero no puede decirnos por qué esta asociación existe y tiene tal nivel. Buscar este porqué es una tarea del investigador, que no puede desentenderse de ella y encomendarla a una fórmula.

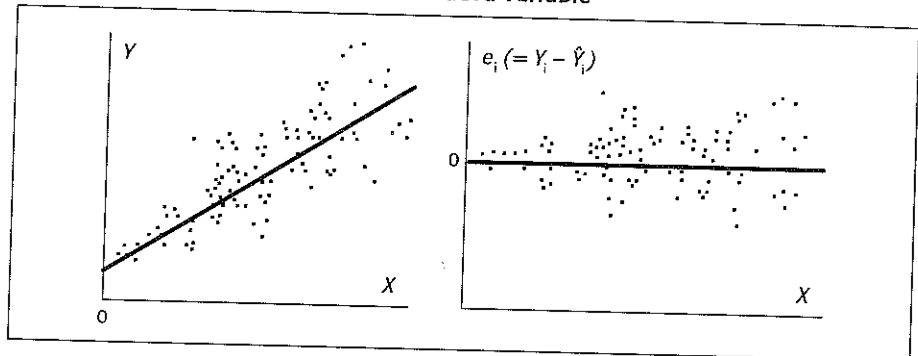
[23] En el capítulo siguiente se tratará el análisis trivariable, mediante tablas, de variables categoriales.

Se puede también dividir cada valor de la variable dependiente Y en una parte reproducida por su asociación con la variable independiente X y una parte residual. Muchos programas de computadora, como el SPSS, producen automáticamente el vector de residuos de una regresión y lo agregan a la matriz original. Si no tenemos estos programas, no es difícil calcular los residuos con esta fórmula:

$$res_{(Y\,en\,X)} = Y_i - a - bX_i.$$

Esta operación matemática corresponde exactamente a la operación gráfica que se presenta en la figura 15.13 y que consiste en una rotación del diagrama tal que pone la recta de regresión paralela a la abscisa.

Figura 15.13: *Cómo los residuos de una regresión se convierten en una nueva variable*



En el diagrama de la derecha, los residuos se relacionan con la misma variable independiente que los produjo. Esto es necesario para mostrar la operación gráfica; pero en lo sustantivo no tendría sentido, porque la recta de regresión no puede ser paralela a la abscisa y por consiguiente *b* y *r* no pueden ser 0. Por otro lado, los residuos pueden ser la variable dependiente en cualquier otro modelo; ellos tienen también una interpretación sustantiva, es decir: la parte de la variabilidad de Y que no puede ser reproducida a través de su asociación con X, y por lo tanto tiene que depender de otras variables no incluidas en el primer modelo bivariable.

CAPÍTULO 16

EL PROCESO DE ANÁLISIS DE TRES VARIABLES CATEGORIALES[1]

Como su nombre lo indica, se denomina análisis trivariable a un diseño de prueba basado en las relaciones entre tres variables. Éste consiste básicamente en complejizar un diseño de dos variables (véase apartado 15.1) a través de la introducción de una tercera, denominada DE PRUEBA O DE CONTROL. La introducción de una tercera variable en una relación bivariada original permite profundizar las hipótesis y enriquecer la explicación de los fenómenos estudiados, dando lugar a mejores aproximaciones a la realidad social donde resulta muy difícil encontrar la explicación al comportamiento de una variable solamente a partir de otra.

El concepto de controlar variables se encuentra en la lógica del método experimental (véase apartado 2.1). Pero dado que este control no es posible en los estudios sociales al modo de los diseños experimentales de las ciencias físico-naturales (véase apartado 2.2), éstos han sido sustituidos por el diseño del cuasi experimento. Este último consiste fundamentalmente en convertir algunas variables en constantes y analizar sus efectos comparando qué ocurre ante su presencia o ausencia.

Si bien se encuentran diseños trivariables en obras clásicas como *El suicidio* de Emile Durkheim (1896), los intentos por formalizar esta estrategia se desarrollaron a mediados del siglo XX cuando se elaboraron las bases lógicas de su justificación. Entre otros, Hyman (1955) encaró el problema de las RELACIONES ESPURIAS (no genuinas) a través de la incorporación de variables de control. De este modo, el autor planteó el esquema del "experimento controlado", a través de la comparación entre las respuestas de un grupo definido como experimental a un estímulo crucial y las de un grupo de control con propiedades equivalentes no expuesto al estímulo. El supuesto que subyace a esta estrategia es que los grupos han sido igualados inicialmente. Es decir que son similares en todas las propiedades excepto el estímulo crucial.

Sin embargo, no es posible garantizar en el cuasi experimento que, al inicio, ambos grupos —experimental y de control— sean equivalentes. Ante la

[1] El modelo de análisis presentado en este capítulo fue desarrollado por Lazarsfeld (1955).

imposibilidad de igualar las unidades de observación en sus condiciones relevantes iniciales, puede ocurrir que la diferencia entre las variaciones observadas de ambos grupos no esté causada por la presencia (ausencia) del estímulo, sino que sea efecto de otras variables que están presentes pero no han sido controladas. En este caso se dice que la relación original encontrada es espuria. La propuesta de Hyman consiste en controlar los posibles terceros factores que podrían invalidar la relación estudiada.

Observemos un ejemplo hipotético, donde + significa alta concentración de frecuencias y – significa bajas frecuencias:[2]

Tabla 16.1: *Educación y voto*

	EDUCACIÓN (x)	
VOTO POR PARTIDOS (y)	POCA	MUCHA
izquierda	+	–
derecha	–	+

En la tabla 16.1 se observa que los más educados tienden a votar por partidos de derecha más que los de menor nivel educativo. Pero si sospechamos que una tercera variable puede estar influyendo en esta relación —por ejemplo 'nivel de ingresos' (variable t)— debemos controlar su efecto a fin de confirmar que la relación xy es genuina. La forma de controlar la tercera variable es convertirla en una constante en nuestro diseño de investigación. Para ello dividimos el total de unidades de la muestra en dos grupos: los de altos y los de bajos ingresos. Y reproducimos la relación original dentro de cada uno de ellos.

Tabla 16.2: *Nivel de ingresos, educación y voto*

	NIVEL DE INGRESOS (t)			
	ALTO (t_1)		BAJO (t_2)	
	EDUCACIÓN (x)		EDUCACIÓN (x)	
VOTO POR PARTIDOS (y)	POCA	MUCHA	POCA	MUCHA
izquierda	–	–	+	+
derecha	+	+	–	–

Hemos convertido la tercera variable (t) en una constante para cada subgrupo de la muestra. En ambas submuestras la relación original (xy) ha desaparecido, lo que nos permite afirmar que esa asociación se debía a la presencia de la tercera variable (t). Es de esperar que ésta presente asociaciones con las

[2] Presentamos todos los ejemplos con variables dicotómicas para facilitar su comprensión.

variables 'educación' (x) y 'voto por partidos' (y). Una configuración posible sería:

Tabla 16.3: *Nivel de ingresos y voto, educación y nivel de ingresos*

	NIVEL DE INGRESOS (t)	
VOTO POR PARTIDOS (y)	BAJO	ALTO
izquierda	+	–
derecha	–	+

	EDUCACIÓN (x)	
NIVEL DE INGRESOS (t)	POCA	MUCHA
bajo	+	–
alto	–	+

La disminución o la desaparición de la relación original dentro de las submuestras, por un lado, y la asociación entre la variable de control (t) y las dos variables de la relación original (x,y), por otro, indican que el factor explicativo de la primera relación encontrada (xy) es la variable t. Es decir, la conducta electoral orientada a elegir partidos de izquierda o derecha no estaría condicionada por la educación sino que se explicaría por el estatus socioeconómico. La tabla 16.3 indica que los individuos de ingresos altos tienden a votar por partidos de derecha y también tienden a ser más educados.

De este modo, el efecto de la variable t, que estaba operando en forma oculta en la relación original, es revelado a través del análisis trivariable.

16.1. Relaciones posibles entre tres variables

Como se pudo observar, al introducir una tercera variable (t) en el análisis de la relación entre otras dos (xy) aparece un conjunto de nuevas relaciones:

1) La relación entre xy para cada uno de los valores de t (t_1, t_2). Es decir: xy para t_1; xy para t_2. Estas relaciones se denominan PARCIALES, ya que son componentes de la relación original.

2) La relación entre la tercera variable (t) y cada una de las variables originales (*x* e *y*). Es decir xt ; ty. Estas relaciones se denominan MARGINALES, pues se construyen a partir de las distribuciones marginales de las relaciones parciales.

Lazarsfeld (1955) presenta un modelo de análisis trivariable que permite elaborar la relación entre dos variables a través de la introducción de una tercera. Con este fin señala que, en un diseño de tres variables, la relación entre *x* e *y*

puede expresarse en una ecuación conformada por las relaciones parciales y marginales (las cuatro relaciones resumidas en los puntos 1 y 2), según la fórmula siguiente:

$$[xy] = [xy, t_1] + [xy, t_2] + [xt] . [ty]$$

Es decir, la relación original es igual a la suma de las dos relaciones parciales más el producto de las relaciones marginales.[3]

Comparando las nuevas relaciones (señaladas en 1 y 2) con la original se pueden presentar tres situaciones:

1) La tercera variable (t) no está significativamente relacionada con las dos variables originales y, por tanto, no afecta al fenómeno en estudio. El factor de prueba es no pertinente[4] y, en consecuencia, las relaciones parciales no difieren de la relación original.

$$[xy] = [xy, t_1] + [xy, t_2] + [0] . [0]$$

2) Se mantiene la relación original y ésta es más fuerte en una de las submuestras, mientras que una de las relaciones marginales es débil. La relación original es elaborada a través de sus relaciones parciales.

$$[xy] = [xy, t_1] + [xy, t_2] + [0] . [ty]$$

En su trabajo sobre el suicidio, Durkheim observa (1896) que entre los católicos se cometen menos suicidios que entre los protestantes y esclarece esta relación incorporando a su análisis la mayor capacidad de integración de la iglesia católica.

3) La variable t está significativamente relacionada con las dos variables originales mientras que las relaciones parciales se debilitan. La relación original es elaborada a través de las relaciones marginales, en la medida que son éstas las que permiten explicar la asociación.

$$[xy] = [0] + [0] + [xt] . [ty]$$

Lazarsfeld (1955) cita el siguiente ejemplo: cuanto más carros de bomberos se llevan a un incendio, mayor es el daño. Al introducir como factor de prueba la magnitud del fuego, los parciales se vuelven cero y las relaciones marginales son fuertes (a mayor fuego, más carros de bomberos y más daño).

[3] Lazarsfeld señala que no se trata de una suma directa sino ponderada.
[4] Son factores pertinentes aquellos que pueden estar asociados a la variable *y*.

Este autor introduce una cuarta variable en el modelo: el orden temporal de aparición de la variable de prueba respecto de la relación original (xy), dando lugar a dos tipos de variable de control:

a) ANTECEDENTE: La variable t es anterior en el tiempo a la relación *xy*. Por ejemplo, si partimos de una relación entre intereses y voto e incluimos como tercera variable al nivel socioeconómico, este último es anterior en el tiempo a las propiedades que componen la relación original. En este caso el nivel socioeconómico es una variable de control antecedente.

b) INTERVINIENTE: La variable t está situada temporalmente entre *x* e *y*. Podemos sostener, siguiendo a Lipset (1959), que el desarrollo económico es un requisito de la democracia, siempre que esté acompañado de legitimidad política. En esta hipótesis la legitimidad política aparece temporalmente como una variable interviniente.

16.2. Los tipos de elaboración según Paul Lazarsfeld

Al combinar los tipos de elaboración (por parciales y marginales) con los tipos de variable de control (antecedente e interviniente), el autor define cuatro tipos posibles de elaboración de una relación original.

TABLA 16.4: Tipos de elaboración de una relación original

	VARIABLE DE CONTROL (t)	
ELABORACIÓN DE XY POR	ANTECEDENTE	INTERVINIENTE
parciales	pa especificación	pi especificación
marginales	ma explicación	mi interpretación

En los casos *pa* y *pi* una de las relaciones parciales es más fuerte que la otra y que la relación original, mientras que las relaciones marginales son más débiles. Es decir, se especifican las circunstancias bajo las cuales la relación original se intensifica y el tipo de elaboración se denomina ESPECIFICACIÓN. Cuando la variable t es antecedente, se ha especificado una condición que intensifica la relación original; cuando t es interviniente se especifica una contingencia.

Ejemplo de especificación
Elaboración por parciales (pi)

Expresaremos las relaciones entre las variables a través del coeficiente de asociación Q.[5]
x = cultura política (más democrática / menos democrática)
y = estabilidad de la democracia (alta / baja)
factor de prueba t = ingreso per cápita (alto/ bajo)

Se observa una relación positiva entre la cultura política de un país y su estabilidad democrática, en el sentido de que cuanto más democrática es la cultura política mayor será su estabilidad democrática ($Q_{xy} = .73$).
Pero esta relación se vuelve más fuerte para los países más ricos ($Q_{xy,t1} = .94$) y tiende a desaparecer en los países más pobres ($Q_{xy,t2} = .10$).
La relación original se ha especificado para un subgrupo de las unidades de análisis.

ESTABILIDAD	CULTURA POLÍTICA (x)	
DEMOCRÁTICA (y)	MÁS DEMOCRÁTICA	MENOS DEMOCRÁTICA
alta	32	9
baja	14	25

$Q_{xy} = .73$

	INGRESO PER CÁPITA ALTO (t_1)		INGRESO PER CÁPITA BAJO (t_2)	
	CULTURA POLÍTICA (x)		CULTURA POLÍTICA (x)	
	MÁS DEM.	MENOS DEM.	MÁS DEM.	MENOS DEM.
alta	22	2	9	8
baja	4	12	11	12

$Q_{xy,t1} = .94 \quad Q_{xy,t2} = .10 \quad Q_{xt} = .48 \quad Q_{ty} = .34$

En los casos *ma* y *mi* las relaciones parciales son más débiles que la relación original y las asociaciones marginales son fuertes. Cuando t es una variable interviniente (situada temporalmente entre *x* e *y*) el tipo de elaboración se denomina INTERPRETACIÓN. Mientras que cuando t es una variable antecedente a la relación original, ésta se explica a través de las relaciones marginales. La importancia de este caso es que permite detectar relaciones originales espurias (no genuinas). Este tipo de elaboración se denomina EXPLICACIÓN.

[5] Q de Yule es una medida de asociación estandarizada para dos dicotomías (véase apartado 15.2), cuyos valores varían desde −1 a +1, expresando estos dos valores la máxima asociación posible y cero la independencia entre las variables. De tal modo que, cuanto más se acerque el coeficiente a cero más débil es la relación entre las variables, y cuanto más se acerque el coeficiente a |1| más fuerte es la relación entre ellas.

Ejemplo de explicación
Elaboración por marginales (ma)

Expresaremos las relaciones entre las variables a través de la diferencia porcentual d%.[6]
x = religión (católica/otra)
y = tipo de matrimonio (religioso/civil)
factor de prueba t = familia de pertenencia (tradicional/moderna)

	RELIGIÓN (x)	
TIPO DE MATRIMONIO (y)	CATÓLICA	OTRA
religioso	84%	11%
civil	16%	89%
N	(107)	(93)

d% = 73

	FAMILIA TRADICIONAL (xy, t_1)		FAMILIA MODERNA (xy,t_2)	
	CATÓLICA	OTRA	CATÓLICA	OTRA
religioso	80%	66%	29%	8%
civil	20%	34%	71%	92%

d% = 14 d% = 21

	RELIGIÓN (x)	
TIPO DE FAMILIA (t)	CATÓLICA	OTRA
tradicional	84%	29%
moderna	16%	71%
N	(107)	(93)

d% = 55

	TIPO DE FAMILIA (t)	
TIPO DE MATRIMONIO (y)	TRADICIONAL	MODERNA
religioso	77%	12%
civil	23%	88%
N	(117)	(83)

d% = 65

[6] La diferencia porcentual (d%) es una medida de asociación estandarizada asimétrica, cuyos valores varían desde −100 a +100, expresando estos dos valores la máxima asociación posible y cero la independencia entre las variables. Se calcula restando las frecuencias relativas en el sentido inverso al que se calcularon los porcentajes. Este coeficiente está basado en la misma lógica que el Q de Yule; se diferencia del mismo por su carácter asimétrico.

Se observa que los católicos contraen matrimonio religioso en mayor medida que los individuos de otras religiones (d% = 73). Sin embargo, al incorporar al análisis el tipo de familia, la fuerte relación original tiende a desaparecer en ambos parciales (d% = 14; d% = 21).
El control de la tercera variable ha puesto de manifiesto que la relación original era espuria y que se explica por la asociación entre el tipo de familia de origen y las dos variables de la relación original. No es la pertenencia a una religión el factor que explica el matrimonio religioso, sino la pertenencia a una familia tradicional.

La importancia del método reside en el control de variables *a posteriori* de la recolección de datos y en la complejización de diseños simples bivariados a través de la incorporación de otras propiedades y asociaciones durante el proceso de investigación.

La lógica del análisis trivariable propuesta por Lazarsfeld (1955) es aplicable a diseños con más de tres variables y a variables con *n* categorías. Sin embargo, cuanto mayor sea el número de variables, mayor el número de sus categorías y más complejo el sistema de asociaciones, más engorrosa será la aplicación del modelo con todos sus pasos. En síntesis, el principal objetivo de la propuesta no es el diseño de un modelo aplicable a una densa red de múltipes relaciones sino que fundamentalmente está orientado a poner en descubierto la lógica que subyace el análisis cuando se trabaja con relaciones más complejas que las bivariadas.

CAPÍTULO 17

OTRAS FORMAS DE ANÁLISIS

17.1. Introducción

Frecuentemente los estudiantes de ciencias sociales manifiestan una clara preferencia por las técnicas de recolección no estándar (observación participante, entrevistas en profundidad, etcétera). Esto se puede deber a muchos motivos: desde la aceptación genuina de sus fundamentos teóricos y epistemológicos, hasta el rechazo —a veces incluso dogmático (véase apartado 1.5)— de todo aquello que se presente como cuantitativo; desde el reconocimiento de técnicas que se basan en gran medida en saberes de la vida cotidiana (observar, conversar, relacionarse socialmente), hasta el desgano por aprender técnicas complejas que juzgan responsables de introducir un tipo de mediación distorsionadora (e inaceptable) entre quien conoce y lo que se conoce. Por otra parte, igualmente frecuente pareciera ser encontrar cierta desazón cuando se trata de "transformar" el producto de tales prácticas de recolección no estándar en algún tipo de objeto tangible —llámese libro, artículo, ponencia, etcétera— en el que se dé cuenta por escrito de los problemas de investigación sobre los cuales se recolectó información —y, cosa no menor, para cuya resolución esta última fue justamente recolectada.

Resulta común encontrar que tales escritos a los que acabamos de hacer referencia no son más que un *collage* de fragmentos de entrevistas o notas observacionales; paradójicamente, el recurso a las técnicas cualitativas, que en la mayoría de los casos podría haber significado un rechazo genuino a las hipersimplificaciones artificiosas resultantes de la investigación estándar (cuantitativa), se revela ahora infantilmente ingenuo: se presentan sus resultados sin más, como si los "datos" hablaran por sí mismos.

En este sentido, resulta oportuno señalar que darles voz a los actores en los relatos formalizados del investigador puede ser un desafío interesante y científicamente valioso; pero esto no se debe entender como mera compilación de sus dichos, sin ningún tipo de mediación y/o interpretación complementaria. Al respecto conviene recordar una de las máximas de la investigación no estándar: las ciencias sociales se caracterizan por las interpretaciones de segundo orden que los científicos hacen a partir de las interpretaciones de primer orden

que los actores verbalizan (por ejemplo, a través de una entrevista) acerca de su vida cotidiana (DOBLE HERMENÉUTICA: Schutz 1962).

Pero estas interpretaciones de segundo orden no son para nada sencillas. La investigación no estándar tiene esta particularidad: sus técnicas de recolección nos resultan —al menos en apariencia— relativamente accesibles; pero sus formas de análisis se tornan escurridizas, intangibles, difíciles de discernir y dominar.

Si todo acto de investigación es una mezcla de conocimientos personales (tácitos) e impersonales, tal vez en ningún otro aspecto, tanto como en el análisis, las formas de investigación estándar y no estándar se encuentren tan en las antípodas. Nada hay más técnico (en el sentido de conocimiento objetivado, impersonal, basado en procedimientos normados, recurrentemente aplicables y ampliamente aceptados: véanse apartados 1.3 y 1.4) en la investigación estándar que sus herramientas de análisis (estadístico); mientras que nada hay menos técnico que el así llamado "análisis cualitativo". Esto no significa que las técnicas estadísticas resuelvan por sí solas los múltiples y complejos aspectos que encierra la mediación entre un conjunto de informaciones empíricamente construidas y algún tipo de interpretación de éstas, tal que permita, eventualmente, objetivar en un texto la descripción o explicación de los fenómenos a los que dichas informaciones remiten. Mucho menos se pretende desconocer las dificultades que los procesos de escritura conllevan. Ya se ha sugerido en otros capítulos cuán relevante es el investigador —en primera persona y sentido fuerte— para seleccionar las técnicas estadísticas adecuadas, para darle sentidos a sus resultados, para relacionarlos con aspectos aparentemente ajenos al confinado mundo sobre el que se ha recolectado información, mucho más allá de lo que un "neutro" coeficiente pueda decir por sí mismo, y para objetivar *ex post* todo el proceso de investigación, y sus resultados, a través de los recursos retóricos que se despliegan en un texto científico.

Más allá de las consideraciones precedentes, lo que resulta claro es que el investigador —sin pretender que sea el intérprete iluminado de alguna supuesta realidad oculta— tiene que decir algo más de lo que las informaciones recolectadas, por sí mismas, podrían llegar a sugerir (si es que esto último fuera en algún modo posible). De lo contrario, su labor pierde sentido, o al menos queda irremediablemente degradada. Y en este decir algo más las técnicas de análisis tienen un papel fundamental: ellas permiten ordenar, sistematizar, preparar la información de una manera funcional a las posibles construcciones discursivas con las cuales se intentará dar cuenta —más o menos convincentemente— de los fenómenos investigados. En este sentido tienen un rol instrumental de mediación y facilitación, y cuando resultan muy esquivas se termina muchas veces en situaciones como las antes descritas: no se llega a construir una interpretación del problema objeto de estudio, a partir de las informaciones recabadas, y sólo se presenta, con más o menos elegancia retórica, una sucesión de tales informaciones "en bruto".

Como ya se ha indicado, esta función de mediación, en el trabajo estándar, la cumplen en gran medida las técnicas estadísticas. Se trata de técnicas en sentido fuerte: a nadie se le ocurre recorrer otros caminos analíticos cuando la base

empírica se expresa (o se puede expresar) a través de números y se condensa (o se puede razonablemente condensar) en una matriz de datos, independientemente de lo que dichos números signifiquen. En todo caso se recorrerán caminos complementarios, pero el recurso a la estadística pareciera ser obligado. En cambio, cuando se trata de bases empíricas constituidas por imágenes y/o palabras, y más aún por textos que no pueden reducirse (o no se quiere reducir) a una suma de palabras, ni tampoco transformar en números de una matriz de datos, el panorama de las "técnicas" de análisis —tal vez en este caso expresión inadecuada— se vuelve mucho más complejo y variado. Por otra parte, los consensos acerca de cuáles sean las más apropiadas en cada caso resultan casi inexistentes. Justamente por estos motivos planteamos que tal vez no sea correcto denominarlas técnicas: no se trata de procedimientos altamente formalizados; no se usan recurrentemente para un mismo tipo de problemas (porque hay instrumentos competitivos/alternativos), y consecuentemente tampoco hay consenso acerca de ellos en la comunidad científica (véase apartado 1.5).

No obstante, con respecto a las así llamadas técnicas de análisis cualitativo (no estándar), hay al menos una cuestión acerca de la cual es relativamente fácil establecer una clasificación útil: nos referimos a las diferentes posiciones sobre la pertinencia (que a veces puede ser simplemente la posibilidad) de su sistematización, dando lugar a técnicas en sentido estricto. Por un lado, están aquellos que creen que el análisis cualitativo es algo artesanal, absolutamente ligado a las capacidades interpretativas del investigador, ya sea que se las conceptualice como un talento innato o como el fruto de aprendizajes (en todo caso a través de procesos no explicitados y difícilmente explicitables). En el otro extremo se ubican quienes creen que el análisis cualitativo se puede basar en técnicas con características análogas a las estadísticas (e incluso algunos proponen procedimientos que permiten reducir las informaciones resultantes del trabajo no estándar a la matriz de datos, y luego tratarlas estadísticamente: véase a continuación apartado 17.2), sólo que, por el supuesto estado de "retraso relativo" en el desarrollo de este tipo de investigación, aún no contamos con tales técnicas.[1] Finalmente, están quienes creen que todo proceso de análisis es complejo, y conlleva múltiples operaciones y decisiones, que a su vez exigen la puesta en juego tanto de conocimientos impersonales (técnicas) como de saberes tácitos y personales, en una infinidad de combinaciones posibles entre ambos (véanse apartados 3.3 y 3.5).

En el caso de la investigación estándar el componente impersonal tiene una presencia relativa mayor; pero esto no significa que en el trabajo no estándar éste sea totalmente inexistente, o que se deba renunciar a cualquier intento de desarrollo de prácticas más o menos formalizadas que sirvan instrumentalmente en nuestros análisis. Por el contrario, tal vez se pueda argumentar que la metodología —tal como ha hecho históricamente— sea en parte responsable

[1] Esta posición resulta cuestionable, especialmente en el sentido de que propone una línea única de desarrollo que las ciencias sociales deberían seguir, y que dentro de las ciencias sociales todas sus variantes deberían seguir, anulando la diversidad —en este caso— en las prácticas de análisis de datos.

de proponer cada vez más soluciones técnicas (conjuntos de conocimientos impersonales) instrumentales a las tareas analíticas e interpretativas, entre otras.[2] En este sentido, repasando la historia de las ciencias sociales se pueden identificar diversas propuestas de sistematización del análisis no estándar. De ellas nos ocuparemos a continuación, aunque sólo en forma panorámica y esquemática.

17.2. El análisis de contenido

Se puede afirmar, en sentido amplio, que el ANÁLISIS DE CONTENIDO "es una técnica de interpretación de textos, ya sean escritos, grabados, pintados, filmados" (Andréu Abela 2001: 2) que se basa en procedimientos de descomposición y clasificación de éstos (Losito 1993). En el marco de una investigación social, los textos de interés pueden ser diversos: transcripciones de entrevistas, protocolos de observación, notas de campo, cartas, fotografías, publicidades televisivas, artículos de diarios y revistas, discursos políticos, etcétera.

Dovring (1954-55: 389), en un artículo sobre la semántica cuantitativa sueca del siglo XVIII, indica que en dicho contexto se dio la primera manifestación del análisis de material impreso de un modo en el que se vislumbran muchos de los problemas centrales del análisis de contenido, tal como se lo entiende en las ciencias sociales en la actualidad. Se trata del estudio del año 1743 sobre una colección de 90 himnos religiosos denominados "Los cantos de Sión", que generó una importante controversia alrededor de lo que hoy podríamos considerar el núcleo del análisis de contenido: ¿quién dijo qué, a quién, cómo y con qué efecto? Pero, como señala Losito (1993), los problemas de método recién se anticipan de modo más evidente en un trabajo de fines del siglo XIX del psicólogo francés Benjamín Bourbon. Con el objetivo de delimitar el campo de la psicología, y diferenciarlo de la filosofía y la religión, Bourbon analizó algunos pasajes del Viejo Testamento, identificando palabras semánticamente medulares —a las que llamó *mots pleines essentiels*— que luego clasificó en categorías de diversos niveles de generalidad.

En el campo más específicamente sociológico suelen destacarse como antecedentes del análisis de contenido las investigaciones de Thomas y Znaniecki (1918-1920) sobre los campesinos polacos, en el marco de las cuales examinaron, entre otros documentos, 754 cartas de inmigrantes de esa nacionalidad en Estados Unidos. Este material fue sometido a un tipo de análisis de contenido no cuantitativo; pero los autores de la Escuela de Chicago no llegaron a explicar sus procedimientos más allá de ciertas cuestiones muy rudimentarias referidas al empleo de una lógica inductiva que permitía evitar conclusiones arbitrarias.

Durante las primeras cuatro décadas del siglo XX, y anticipando en cierto sentido los actuales *media studies* (al menos en algunas de sus corrientes y líneas de investigación), comenzó a difundirse en Estados Unidos el interés por analizar materiales periodísticos, alcanzando un gran nivel de desarrollo y diversificación. Dichas investigaciones se llevaban a cabo fundamentalmente en la Universidad de Columbia, y derivaban del reconocimiento de que los medios jugaban un papel político, social y cultural cada vez más importante en la sociedad de la época. Esta cuestión ya había sido destacada por Max Weber en un congreso de 1910 en Alemania, quien incluso propuso investigaciones empíricas sobre los medios de prensa, que por distintos motivos no llegaron a realizarse (Losito 1993).

[2] Sensibilizando a su vez acerca del innegable (e importantísimo) papel que tiene el conocimiento personal en dichas tareas de análisis e interpretación.

Pero la formalización metodológica del análisis de contenido sólo empieza a tomar forma en los años cuarenta y cincuenta del siglo pasado, cuando se publican los primeros manuales especializados. La obra de Berelson (1952), *Content Analysis in Comunication Research*, que retoma desarrollos previos del mismo autor junto con Laszarsfeld, se convertiría casi inmediatamente en la referencia obligada. En su manual, Berelson (1952: 18) define al análisis de contenido como "una técnica de investigación para la descripción objetiva, sistemática y cuantitativa del contenido manifiesto de la comunicación". El énfasis puesto en la objetividad deja entrever que se trata de una perspectiva inscrita en el marco de la concepción estándar de la ciencia, y más específicamente, del consenso ortodoxo típico de las ciencias sociales norteamericanas de la época (véase apartado 2.2).

Por otra parte, tal como indica Stemler (2001), el interés por el tratamiento cuantitativo de los aspectos manifiestos del texto se basa en el presupuesto de que las palabras más mencionadas en él permiten dar cuenta del verdadero foco de una comunicación. Sin embargo, pronto se reconoció que el análisis de contenido debería ir más allá de un simple conteo de palabras, y en este sentido la propuesta de Berelson fue cuestionada, dando lugar a otras formas de análisis de contenido que incorporaron la exploración de cuestiones latentes, o al menos —como sugiere Krippendorff (1980)— del contexto, en cuanto "marco de referencia donde se desarrollan los mensajes y los significados" (Andréu Abela 2001: 3).

En la actualidad se suele distinguir entre el análisis de contenido clásico, fiel a la cuantificación de los aspectos manifiestos del texto (aunque en versiones que incorporan algunos elementos contextuales), y el análisis de contenido cualitativo, desarrollado a partir de los años ochenta, que se basa en un conjunto de técnicas destinadas a interpretar su sentido latente u oculto.

El reconocimiento de las consistentes críticas que se le hicieron durante los últimos cincuenta años, así como la consecuente incorporación de elementos "correctivos" en los métodos de análisis de contenido, hace que estudios de tipo clásico idénticos a los propuestos en la década de 1950 sean en el presente una absoluta rareza. En cambio, como señala Andréu Abela (2001: 20-22), el análisis de contenido clásico se presenta ahora a través de tres tipos: el análisis temático, el análisis semántico y el análisis de redes. El análisis de contenido temático se centra en la presencia de términos, con independencia de las relaciones que surjan entre ellos; "las técnicas más utilizadas son las listas de frecuencias, la identificación y clasificación temática, y la búsqueda de palabras en

contexto". El análisis de contenido semántico pretende ante todo estudiar las relaciones entre los temas tratados en un texto; para esto se define cierta estructura significativa de relación y se consideran todas las ocurrencias que concuerden con ella. Finalmente, el análisis de contenido de redes se centra en la ubicación relativa de ciertos componentes en el texto, asumiendo que la "red léxica ideológicamente significativa que impregna el discurso se identifica sobre la base de la reiteración y da como resultado un esquema de la organización semántica de este discurso en forma de red".

El análisis de contenido cualitativo no niega las ventajas del cuantitativo, al que suele tomar como una etapa inicial; sólo que busca enriquecerlo a través de procedimientos interpretativos, tratando de ir más allá de los aspectos manifiestos a través de la consideración del contenido latente y del contexto en el que se inscribe un determinado texto (Andréu Abela 2001).

Krippendorff (1980) sugiere que al hacer análisis de contenido el investigador debe comenzar por plantearse una serie de preguntas:

1) ¿Qué datos serán analizados?
2) ¿Cómo se los define?
3) ¿Cuál es la población de la cual se los toma?
4) ¿Cuál es el contexto relativo a los datos analizados?
5) ¿Cuáles son los límites del análisis?
6) ¿Cuál es el objetivo de las inferencias?

Sin embargo, debe quedar claro que estas preguntas, en conjunto, tienen sentido si el análisis de contenido se refiere a la estrategia global de investigación, no si se recurre a él instrumentalmente con el único fin de analizar textos producidos en el marco de otro tipo de estudio (por ejemplo cuando su uso se limita al análisis de preguntas abiertas en una encuesta, o a las transcripciones de entrevistas en profundidad).

En todo caso, una vez definido el conjunto de textos (o imágenes a analizar), la cuestión fundamental es construir un esquema de codificación (*coding frame*). Esto se puede hacer *a priori*, cuando las categorías que orientarán el análisis se definen de antemano, a partir de una opción teórica, o siguiendo un recorrido emergente. Esta última modalidad se encuentra mucho más difundida, y consiste en la construcción de una suerte de cuestionario estandarizado *ex post* a partir de una exploración preliminar de los textos. En efecto, la lectura de algunos exponentes del conjunto de materiales por analizar permite identificar los temas abordados en éstos, así como las categorías relativas a ellos que aparecen en el texto. En otras palabras, se trata de identificar las "respuestas implícitas" en el texto a un cuestionario imaginario al cual no responde, pero al que podría haber respondido. El esquema de codificación incluirá todas estas preguntas imaginarias (que pueden ser sólo enunciaciones de temas) y las distintas respuestas que para ellas aparecen en los textos preexaminados. Se exploran varios textos en esta etapa de construcción del esquema justamente para darle un carácter más completo: basarse en un solo texto podría provocar importantes sesgos, o en todo caso una notable incompletitud.

Con respecto a la construcción del esquema se ha señalado críticamente la tendencia de los investigadores a la autogratificación (*self-fulfilment*), es decir, la tendencia a "imponer" categorías que se ajusten a sus hipótesis y presupuestos, sean explícitos o no. Asimismo, se ha cuestionado el énfasis puesto en aquello que en el texto aparece de modo manifiesto. Para atender a la primera objeción se ha propuesto que la construcción del esquema se haga a partir del consenso entre codificadores (*inter-coder agreement*). Esto implica dos cosas: 1) que distintas personas construyan marcos individualmente, y que luego alcancen consensos contrastando las propuestas de cada una de ellas, y 2) que una vez definida una versión final del esquema de codificación, ella sea aplicada uniformemente por todos los codificadores (en caso que haya más de uno), considerando de igual modo cada pasaje del texto analizado. Con respecto a la segunda objeción, se ha propuesto incluir en el esquema aspectos contextuales no explícitos, referidos tanto a quien produjo el texto (por ejemplo variables sociodemográficas, afiliación política, etcétera) como al texto mismo (extensión, tipo de relato, fecha de producción, etcétera).

Es evidente que la primera de las objeciones se relaciona directamente con los clásicos problemas de fiabilidad. Al respecto, el acuerdo entre codificadores también se suele tomar como base de los más difundidos y tradicionales exámenes de fiabilidad en el marco del análisis de contenido. Para esto se recurre generalmente a un coeficiente llamado K de Cohen:

$$K = \frac{P_A - P_c}{1 - P_c}$$

donde P_A es la proporción de unidades en la que los codificadores acuerdan y P_c es la proporción de unidades en las que el acuerdo se esperaría sólo por azar. Para calcular este coeficiente se parte de una matriz de acuerdos (una simple tabla de contingencia) en la que cada codificador (de una pareja) representa respectivamente las filas y las columnas. En las celdas se ubican tanto las frecuencias empíricas (observadas) como las teóricas (esperadas) (véase apartado 15.2). La sumatoria de las diagonales definidas desde la primera celda (arriba a la izquierda) hasta la última (abajo a la derecha), tanto de las frecuencias observadas como de las esperadas, resulta en P_A y P_c respectivamente. El coeficiente adquiere el valor 1 cuando el acuerdo en la codificación es perfecto, y 0 cuando no hay más acuerdo que el esperado al azar. En general, se suelen tomar como sustanciales valores del coeficiente mayores a .6 (Stemler 2001).

Una vez alcanzada la versión final del esquema de codificación, y habiendo controlado el funcionamiento efectivo de los acuerdos entre codificadores en las pruebas piloto, éste se aplica a todos los textos en cuestión. Se trata de interpelarlos de manera estandarizada a partir de este esquema de codificación. Por otra parte, el esquema constituye la base estructural de una matriz de datos, que luego será completada con los resultados de su aplicación al conjunto de textos bajo análisis. Queda claro que una vez construida y completada la matriz se procede a su análisis estadístico, recurriendo a las técnicas univariadas, bivariadas y multivariadas habituales (véanse capítulos 14, 15 y 16).

Sin embargo, el recurso a las técnicas tradicionales —por vía del análisis de contenido clásico— no constituye la única manera de abordar estadísticamente un texto. Aunque de difusión más limitada, también es frecuente seguir los modelos lexicométricos desarrollados en el marco de la escuela francesa de estadística textual (*Analyse des Données*: Benzécri 1973, Bécue 1991, Lebart y Salem 1994). Se trata de enfoques más habituales en los estudios de comunicación, opinión pública y *marketing*. Ellos se basan en el conteo de las frecuencias de unidades verbales básicas (en general las palabras) sin que medien codificaciones previas, sobre las que luego se realizan distintos tipos de análisis estadísticos (especialmente análisis de correspondencias). Para esto se dispone de *softwares* específicos, especialmente el SPAD.T (*Système Portable pour l'Analyse des Donées Textuelles*).

17.3. La teoría fundamentada y el método comparativo constante

En el capítulo 2 se ha señalado que a partir de la década de 1960 se inició una etapa de importante revitalización de los enfoques de investigación no estándar. Este nuevo auge trajo aparejados debates acerca del estatus científico de éstos. Como ya se ha visto en dicho capítulo, en aquella época —y en parte aún en la actualidad— se los cuestionaba por sus inherentes problemas de fiabilidad y por las restringidas posibilidades de generalización de sus resultados. Por otra parte, esto se pensaba como una limitación indisolublemente ligada a la falta de estandarización de sus procedimientos de investigación. Así las cosas, comenzó a difundirse una visión que, aun aceptando los métodos no estándar como parte legítima del patrimonio de las ciencias sociales, los confinaba a un rol secundario, útiles sólo para las primeras exploraciones de fenómenos poco conocidos. Debe quedar claro que tal posición era significativa especialmente en la sociología; en la antropología, en cambio, en la que ya desde los años veinte la observación de campo constituía una práctica canónica —un *rite de passage* como la llama Burgess (1984)—, la situación era un poco diversa. Las críticas a los métodos no estándar también se habían difundido en esta última disciplina, y en este sentido se debe tener en cuenta que ya en la década de 1960 el trabajo etnográfico se aplicaba ampliamente al estudio de fenómenos de la sociedad industrial (que en la distinción tradicional eran objeto de la sociología y no de la antropología, más centrada en el estudio de las sociedades "primitivas" y tradicionales). Pero la antropología, ámbito de desarrollo de corrientes intelectuales ligadas al relativismo descriptivo y al particularismo, no se interesó tanto como la sociología por sistematizar procedimientos impersonales que dieran remedio a una supuesta falta de cientificidad, entendida esta última en los términos clásicos de las ciencias físicas. En la sociología —disciplina muchas veces en tensión (por emulación o rechazo) con tales ciencias (véanse capítulos 1 y 2), y a su vez más permeada por corrientes comprometidas con las explicaciones globales— aun muchos de aquellos empeñados en la defensa de su especificidad trataron de dar respuesta a este problema promoviendo algún tipo de estandarización de los procedimientos. El aspecto más delicado en este sentido siempre fue el relativo al análisis.

Tal vez el primer gran intento orientado a su formalización haya sido el de Znaniecki, que en un su clásico *The Method of Sociology* (1934) propuso a la inducción analítica como un procedimiento sistemático para desarrollar clasificaciones y tipos, o para producir hipótesis explicativas a partir del trabajo cualitativo (Valles 1997). Pero la propuesta de Znaniecki fue severamente criticada, especialmente por su pretensión de formular leyes universales y, junto con la decadencia de los métodos no estándar luego de la Segunda Guerra Mundial (véase apartado 2.3), sus aplicaciones perdieron vigencia.

Distinta ha sido la suerte que corrió el segundo gran intento de proveer pautas sistemáticas para el análisis de la información en el marco de los estudios de caso. Nos referimos a la propuesta de Glaser y Strauss (1967) acerca de los procedimientos para generar teoría fundamentada (*grounded theory*). Aunque no exento de críticas, su modelo, para cuya aplicación incluso existen en la actualidad conocidos *softwares* (por ejemplo el programa alemán ATLAS.ti), sentó las bases de mucho de lo que hoy podría considerarse práctica común en el análisis cualitativo de transcripciones de entrevistas y otros materiales afines.

A fin de generar teoría de alcance medio, Glaser y Strauss (1967) presentaron un procedimiento que llamaron MÉTODO COMPARATIVO CONSTANTE, en cierto sentido opuesto a la estrategia canónica de generación de teoría a partir de deducciones. Según su formulación original, el procedimiento consta de cuatro etapas, desde la comparación de incidentes (fragmentos de datos derivados de observaciones y entrevistas - *slices of data*) hasta la escritura de la teoría, pasando por la integración de categorías y sus propiedades, y la delimitación de la teoría (Valles 1997).

El primer paso consiste en comparar distintos fragmentos de datos, seleccionados a partir del criterio de muestreo teórico, y agrupar todos aquellos que remitan a una misma cuestión bajo un rótulo común, hasta ir "descubriendo" categorías conceptuales, y sus dimensiones, que deberán ser saturadas (véanse apartados 1.3 y 12.5). Luego, las categorías y dimensiones ya saturadas (además de las hipótesis que surgen durante este proceso de codificación, relativas a interpretaciones preliminares acerca de las relaciones entre las categorías) se van articulando e integrando. En este proceso de delimitación de la teoría se buscará, por un lado, satisfacer el criterio de parsimonia (*parsimony*) a través de la selección de aquellas categorías —y sus (inter)relaciones— que sean centrales (con la consecuente transformación, integración o eliminación de las categorías originales). Por otro lado, se determinarán las posibilidades de generalización de la teoría más allá del conjunto de informaciones sobre el caso específico que le dieron lugar. En todo este recorrido se recurre a la comparación constante, una suerte de diálogo entre teoría y base empírica (aunque tomando como punto de partida las informaciones derivadas del trabajo de campo).

Como se ha dicho antes, la propuesta de Glaser y Strauss ha sido objeto de críticas. Se han señalado especialmente los riesgos de inductivismo; el excesivo énfasis puesto en la generación de teoría, a expensas de la captación de la experiencia vivida en las situaciones de interacción, y la carencia de principios suficientemente concretos sobre cómo implementar en la práctica el método comparativo constante (Valles 1997). Sin embargo, y a pesar de que no siempre sea

aplicada en su formulación original, la gran difusión de la teoría fundamentada en la ciencias sociales —como se dijo más arriba— permitió sentar algunas bases elementales, y en la actualidad relativamente consensuadas, acerca de cómo encarar el análisis de las entrevistas. En efecto, la identificación de núcleos temáticos y su consecuente articulación, determinando los mapas conceptuales presentes en el discurso —aunque en cierto sentido elemental— sigue siendo práctica común, especialmente dada la complejidad de las técnicas alternativas de análisis del discurso que descansan sobre conocimientos lingüísticos y semióticos fuera del alcance de la mayoría de los investigadores sociales, y que abordaremos a continuación.

17.4. El análisis del discurso

La expresión 'análisis del discurso' es muy reciente; comenzó a utilizarse en inglés (*discourse analysis*) a partir de su aparición en el título de un artículo del lingüista norteamericano Zellig Harris de principios de la década de 1950. Sin embargo, en sus perspectivas actuales, que se nutren de elementos de la semiótica, la psicolingüística, la sociolingüística, la pragmática, la etnografía del habla, etcétera, empezó a desarrollarse entre fines de los años sesenta y principios de los setenta.

Con esta expresión se alude a un conjunto amplio de enfoques cuyo objetivo central es el análisis del lenguaje hablado y escrito. Para Van Dijk (1988), se trata de una práctica caracterizada por la transdisciplinariedad, el interés por el texto y el contexto, por el habla en la vida cotidiana y por la variedad de géneros discursivos. Vasilachis (1992: 176-177) indica que "el análisis del discurso tiende a ser una práctica académica interdisciplinaria y socialmente orientada", cuyo "núcleo clásico son los análisis estructurales de las propiedades lingüísticas, estilísticas, retóricas o pragmáticas de los textos y del habla".

Las corrientes que se sitúan bajo el rótulo de análisis del discurso son muy diversas y sus aplicaciones se manifiestan en diferentes campos del conocimiento humano; éstas exceden ampliamente el tratamiento que se le dará al tema en este capítulo, que constituye tan sólo una mirada panorámica. Por otra parte, el análisis del discurso se puede considerar un tipo de investigación no estándar en sí mismo, y no sólo una estrategia o mero procedimiento de análisis de la información cualitativa. Por esta razón, tal como afirma Alonso (1998), la delimitación del lugar práctico que éste ocupa en la sociología representa un escollo importante.

Para tratar de clarificar la cuestión, este mismo autor distingue genéricamente "tres niveles básicos de aproximación al análisis de discursos: un nivel informacional/cuantitativo, un nivel estructural/textual y un nivel social-/hermenéutico" (*ibid.*: 189).

El primero de ellos se centra en la exploración de las dimensiones denotativas y manifiestas del texto, y para esto recurre al tratamiento cuantitativo/estadístico. Obviamente, nos estamos refiriendo a lo que hemos denominado análisis de contenido y que ya hemos tratado en el apartado 17.2.

El análisis estructural surge para Alonso (*ibid.*: 195) del entrecruzamiento entre la lingüística estructural y la antropología, y busca "hallar los elementos lógicos invariantes y prácticamente universales de los textos concretos, o [...] las estructuras *subyacentes* abstractas que tienden a cumplirse también en este o aquel texto elegido". El estructuralismo, originado en la lingüística de Saussure y caracterizado por concebir a la realidad como construida por el lenguaje, se difundió en las ciencias sociales empíricas de la mano del antropólogo francés Lévi-Strauss. Él tomó a la lingüística como modelo para sus estudios de la sociedad —en gran medida gracias a su trabajo conjunto con el lingüista ruso Jakobson. Por otra parte, la corriente estructuralista ha sido central en lo que se conoce como giro lingüístico, en cuanto considera al lenguaje como agente estructurante. Además de Lévi-Strauss, también se han destacado como analistas estructurales del relato Greimas y Barthes.

A pesar de su importancia y éxito académico, Alonso (*ibid.*: 202-204) destaca los límites de este enfoque para la investigación social empírica; a su juicio el sociólogo no debe centrarse en el texto por sí mismo sino en cuanto soporte de discursos, desde una perspectiva sociohermenéutica: "nos interesa lo que los discursos hacen —y de lo que se hacen los discursos— y no lo que los textos formalizan desde su nivel genético, lo que implica una visión *pragmática* del discurso, pero no tanto una pragmática lingüística de la microsituación intercomunicativa, como una macropragmática referida a los espacios y conflictos sociales que producen, y son producidos, por los discursos". En este sentido, es fundamental situar al texto en contexto, pero no sólo entendido como marco situacional de la producción del discurso, o como contexto semiótico/intertextual, sino como contexto social ("conjunto de referentes sociogrupales que otorgan significaciones al discurso").

Las consideraciones precedentes refuerzan la idea expresada al inicio de este apartado: la noción de análisis del discurso es ambigua y engloba una gran variedad de perspectivas. En el marco de las variantes que Alonso (1998) define como sociohermenéuticas, podría tomarse como aproximativa la definición de Stubbs (1983: 1): análisis del discurso se refiere en general a los intentos por estudiar la organización del lenguaje más allá de la oración, y que se centran en el análisis de éste en su contexto social.

De todas las corrientes vinculadas a esta definición amplia, tal vez la más influyente en las ciencias sociales empíricas contemporáneas sea la que se conoce como análisis crítico del discurso, una perspectiva interdisciplinaria que comenzó a desarrollarse a partir de los años ochenta. Como sostiene Fairclough (1989), ésta considera al lenguaje en cuanto forma de práctica social, centrándose en los modos en que la dominación política y social se reproduce a través de los textos. Además de Fairclough, otros destacados exponentes de esta corriente son Wodak, Van Dijk y Van Leeuwen. Más allá de sus aportes teórico-metodológicos (Fairclough 1989, 1995, 2003; Van Dijk 1977, 1978; Van Leeuwen 2005), los cultores del análisis crítico del discurso son reconocidos por sus investigaciones sobre género (Wodak 1997), racismo (Van Dijk 1991, 2003) e ideología (1998).

En el contexto latinoamericano, Vasilachis (1992, 1997) ha sido una de las pioneras en aplicar elementos del análisis del discurso a la investigación social

empírica. Su propuesta metodológica pretende abordar "el espacio en que convergen los métodos cualitativos y el análisis lingüístico de textos orales o escritos, de acuerdo con la hipótesis de que ambas estrategias comparten [...] supuestos del paradigma interpretativo" (1992: 153). En sus desarrollos confluyen elementos diversos: la teoría de la estructuración de Giddens, la teoría de la acción comunicativa de Habermas, y las teorías de los actos del habla de Austin, Searle y Habermas. Su propuesta se articula a partir de cuatro tesis fundamentales:

1) el lenguaje es una forma de producción y reproducción del mundo social;
2) el contexto determina el significado y alcance de las emisiones y su producción, así como el contenido de las interpretaciones;
3) el análisis lingüístico puede contribuir a la comprensión del sentido de la acción en la interacción comunicativa, y
4) el intérprete debe asumir la perspectiva de los actores y reconocer los significados que ellos le dan a su acción y a la de otros dentro de un contexto dado.

Sin embargo, más allá de todos los importantes desarrollos que se han dado en esta materia en las últimas décadas, aún no se asiste —en el ámbito de las ciencias sociales empíricas— a una universalización de sus aplicaciones. Tres son las cuestiones que se han de tener en cuenta en este sentido. Por un lado, se debe considerar el hecho de que no existe una estrategia única, de consenso (o al menos una que se haya erigido como dominante), sobre cómo abordar el análisis del texto en contexto. Por otro lado, se debe reconocer que muchas de las propuestas son fundamentalmente formas de investigación no estándar en sí (con sus supuestos teóricos —especialmente sobre el sujeto, el lenguaje y la comunicación— epistemológicos y metodológicos); no se trata de meros conjuntos de técnicas que puedan fácilmente reducirse a usos tan sólo instrumentales en el marco de otro tipo de investigaciones. Finalmente, cabe destacar que la mayoría de ellas —como ya se ha dicho— exigen conocimientos sofisticados (de lingüística y semiótica, entre otros) que no están al alcance de la mayoría de los investigadores sociales, y que en todo caso —y a diferencia de las técnicas estadísticas de análisis— rara vez forman parte de su *curricula* de estudio.

Bibliografía

ADLER, Patricia A. y Peter ADLER (1994) *Observational Techniques*, pp. 377-392 en Norman K. Denzin e Yvonna S. Lincoln (eds.), *Handbook of Qualitative Research*. Thousand Oaks: Sage.

ADORNO, Theodor Wiesengrund *et al.* (1950) *The Authoritarian Personality*. New York: Harper & Row.

AGNOLI, Maria Stella (1992) *La costruzione delle variabili tra rilevazione e analisi dei dati: il caso dei sondaggi*, pp. 142-163 en Alberto Marradi y Giancarlo Gasperoni (comps.), *Costruire il dato 2*. Milano: Franco Angeli.

ALDENDERFER, M. S. y R. K. BLASHFIELD (1985) *Cluster Analysis*. London: Sage.

ALLEN, Michael Patrick (1974) *Construction of Composite Measures by the Canonical-Factor-Regression Method*, pp. 51-78 en Herbert L. Costner (comp.), *Sociological Methodology 1973-74*. San Francisco: Jossey Bass.

ALONSO, Luis E. (1998) *La mirada cualitativa en sociología*. Madrid: Fundamentos.

AMISANO, Ernesto, Giuseppe RINALDI y Emilia PAMPANIN (2002) *Sono equidistanti le categorie di una scala Likert?*, pp. 95-122 en Marradi y Gasperoni (2002).

AMMASSARI, Paolo (1985) *I fondamentali problemi di metodologia della ricerca sociale*, en "Studi di Sociologia" XXIII, 2-3 (aprile): 176-93.

ANASTASI, Anne (1953) *Differential Psychology*. London: Macmillan.

ANASTASI, Anne (1954) *Psychological Testing*. London: Macmillan.

ANDRÉU ABELA, Jaime (2001) *Las técnicas de análisis de contenido: una revisión actualizada*, en "Documento de trabajo CENTRA" 2001/03, <http://public.centrodeestudiosandaluces.es/pdfs/S200103.pdf>.

ARMOR, David J. (1974) *Theta Reliability and Factor Scaling*, pp. 17-50 en Herbert L. Costner (comp.), *Sociological Methodology 1973-1974*. San Francisco: Jossey-Bass.

ARDÈVOL, Elisenda *et al.* (2003) *Etnografía virtualizada: la observación participante y la entrevista semiestructurada en línea*, en "Athenea Digital" n. 3: 72-92.

ARNAULD, Antoine y Pierre NICOLE (1662) *Logique, ou art de penser*. Paris.

ARON, Raymond (1996) *Lecciones sobre la historia. Cursos del Collège de France*. México: FCE.

BACHELARD, Gaston (1934) *Le nouvel esprit scientifique*. Paris: Alcan.

BACHELARD, Gaston (1938) *La formation de l'esprit scientifique. Contribution à une psychanalyse de la connaissance objective.* Paris: Vrin.
BACON, Sir Francis (1620) *Novum Organum.*
BACHMANN, Duane, John ELFRINK y Gary VAZZANA (1996) *Tracking the Progress of E-Mail vs. Snail-Mail* en "Marketing Research" VIII, 2 (summer): 30-6.
BAILYN, Lotte (1977) *Research as a Cognitive Process: Implications for Data Analysis,* en "Quality and Quantity" XI, 2 (june): 97-117.
BARTLETT, Claude J. *et al.* (1960) *A Comparison of Two Methods of Attitude Measurement: Likert Type and Forced Choice,* en "Educational and Psychological Measurement" XX, 4 (winter): 699-704.
BECKER, Howard S. (1970) *Sociological Work: Method and Substance.* Chicago: Aldine.
BÉCUE, Mónica (1991) *Análisis de datos textuales. Métodos estadísticos y algoritmos.* París: CISIA.
BELSON, William A. (1981) *The Design and Understanding of Questions in the Survey Interview.* London: Gower.
BENZÉCRI, Jean-Paul (1973) *L'analyse des données.* Paris: Dunod.
BERELSON, Bernard (1952) *Content Analysis in Communication Research.* New York: Free Press.
BERG, Irwin A. (comp., 1967) *Response Set in Personality Assessment,* Chicago: Aldine.
BERKA, Karel (1983) *Measurement. Its Concepts, Theories and Problems.* Dordrecht: Reidel.
BERKOWITZ, Norman H. y George H. WOLKON (1964) *A Forced Choice Form of the F Scale - Free of Acquiescence Response Set,* en "Sociometry" XXVII, 1 (march): 54-65.
BERNARD, H. Russell (1994) *Research Methods in Anthropology: Qualitative and Quantitative Approaches.* Thousand Oaks: Sage.
BISHOP, George F. *et al.* (1982) *Effects of Presenting One versus Two Sides of an Issue in Survey Questions,* en "Public Opinion Quarterly" XLVI: 69-85.
BISHOP, George F. *et al.* (1989) *Psuedo-Opinions on Public Affairs,* pp. 425-36 en Eleanor Singer y Stanley Presser (comps.), *Survey Research Methods.* University of Chicago Press.
BLALOCK, Hubert M. (1961) *Causal Inferences in Nonexperimental Research.* Chapel Hill: University of North Carolina Press.
BLALOCK, Hubert M. (1969) *Theory Construction. From Verbal to Mathematical Formulations.* Englewood Cliffs: Prentice-Hall.
BLOCK, Jack (1965) *The Challenge of Response Sets.* New York: Appleton-Century-Crofts.
BLUMER, Herbert (1969) *Symbolic Interactionism: Perspective and Methods.* Englewood Cliffs: Prentice-Hall.
BOCCUZZI, Ermelinda (1985) *Parcellizzazioni e reificazioni nelle ricerche sociologiche: il punto di vista di un'intervistatrice,* en "Rassegna Italiana di Sociologia" XXVI, 2: 239-260.
BOHRNSTEDT, George W. (1970) *Reliability and Validity Assessment in Attitude Measurement,* pp. 80-99 en Gene F. Summers (comp.), *Attitude Measurement.* Chicago: Rand-McNally.

BOURDIEU, Pierre (1973) *L'opinion publique n'existe pas,* en "Les temps modernes" n. 318 (janvier): 1292-1309.
BOURDIEU, Pierre (1979) *La distinction. Critique sociale du jugement.* Paris: Ed. de Minuit.
BOURDIEU, Pierre (1994) *Raisons pratiques. Sur la théorie de l'action.* Paris: Seuil.
BOURDIEU, Pierre, Jean-Claude CHAMBOREDON y Jean-Claude PASSERON (1968) *Le métier de sociologue. Préalables épistémologiques.* Paris: Mouton.
BRAITHWAITE, Richard (1953) *Scientific Explanation: A Study of the Function of Theory, Probability and Law in Science.* Cambridge University Press [citas de la edición 1960].
BRENNER, Michael *et al.* (1985) *The Research Interview: Uses and Approaches.* London: Academic Press.
BRISLIN, Richard W., Walter J. LONNER y Robert M. THORNDIKE (1973) *Cross-Cultural Research Methods.* New York: Wiley.
BROWNELL, William A. (1933) *On the Accuracy with which Reliability May Be Measured by Correlating Test Halves,* en "Journal of Experimental Education" I: 204-15.
BRUSCHI, Alessandro (1991) *Logica e metodologia,* en "Sociologia e Ricerca Sociale" XII, 35: 30-55.
BRYMAN, Alan E. (1988) *Quantity and Quality in Social Research.* London: Routledge.
BRYMAN, Alan E. (2004) *Triangulation,* en Michael S. Lewis-Beck *et al.* (eds.), *Encyclopedia of Social Science Research Methods.* London: Sage.
BUCHANAN, William y Hadley CANTRIL (1953) *How Nations See Each Other: A Study in Public Opinion.* Urbana: Illinois University Press.
BUNGE, Mario A. (1973) *On Confusing 'Measure' with 'Measurement' in the Methodology of the Behavioral Sciences,* en Bunge (comp.), *The Methodological Unity of Science.* Dordrecht: Reidel.
BURGESS, Robert G. (1984) *In the Field. An Introduction to Field Research.* London: Routledge.
CACCIOLA, Salvatore y Alberto MARRADI (1988) *Contributo al dibattito sulle scale Likert basato sull'analisi di interviste registrate,* pp. 63-105 en Marradi (comp.), *Costruire il dato. Sulle tecniche di raccolta delle informazioni nelle scienze sociali.* Milano: Franco Angeli.
CALVI, Gabriele (1980) *La classe fortezza. Scelte degli elettori e responsabilità della classe politica in Italia.* Milano: Angeli.
CAMPBELL, Donald T. y Donald Winslow FISKE (1959) *Convergent and Discrimination Validation by the Multitrait-Multimethod Matrix,* en "Psychological Bullettin" LVI, 2 (march): 81-105.
CAMPBELL, Donald T. y Julian C. STANLEY (1963) *Experimental and Quasi-Experimental Designs for Research.* Chicago: Rand-McNally [tr. esp. *Diseños experimentales y cuasi experimentales en la investigación social.* Buenos Aires: Amorrortu 1976].
CAMPBELL, Norman R. (1928) *An Account of the Principles of Measurement and Calculation.* London: Longmans & Green.
CAMPELLI, Enzo (1991) *Il metodo e il suo contrario.* Milano: Angeli.

CANTRIL, Hadley y Lloyd A. FREE (1962) *Hopes and Fears for Self and Country: The Self-Anchoring Striving Scale in Social Research*, en "American Behavioral Scientist" VI, 2 (october): supplement.
CAPECCHI, Vittorio (1972) *Struttura e tecniche della ricerca*, pp. 23-120 en Pietro Rossi (comp.), *Ricerca sociologica e ruolo del sociologo*. Bologna: Il Mulino.
CARDANO, Mario (1991) *Il sociologo e le sue muse: qualità e quantità nella ricerca sociologica*, en "Rassegna italiana di sociologia" XXXII, 2: 181-223.
CARDOSO DE OLIVEIRA, Roberto (1996) *O trabalho do antropologo: olhar, ouvir, escrever*, en "Revista de Antropologia" XXXIX, 1: 13-37.
CARMINES, Edward G. y Richard A. ZELLER (1979) *Reliability and Validity Assessment*. London: Sage.
CARNAP, Rudolf (1939) *Foundations of Logic and Mathematics*. Chicago: University Press.
CARNAP, Rudolf (1950) *Logical Foundations of Probability*. Chicago: University Press.
CHIESI, Leonardo (2002) *Qualità e quantità: un* outline *del dibattito e una proposta*, en "Sociologia e ricerca sociale" n. 67: 180-192.
CICOUREL, Aaron Victor (1964) *Method and Measurement in Sociology*. New York: Free Press.
CINI, Marcello (1990) *Trentatré variazioni sul tema. Soggetti dentro e fuori la scienza*. Roma: Editori Riuniti.
CLARK, Herbert y Michael SCHOBER (1991) *Asking Questions and Influencing Answers*, pp.15-43 en Judith M. Tanur (comp.), *Questions About Questions: Inquiries Into the Cognitive Bases of Surveys*. New York: Russell Sage Foundation.
COHEN, Morris Raphael y Ernst NAGEL (1934) *An Introduction to Logic and Scientific Method*. New York: Harcourt.
COLLINS, Randall (1975) *Conflict Sociology: Toward an Explanatory Science*. New York: Academic Press.
COMBESSIE, Jean-Claude (1996) *La méthode en sociologie*. Paris: La Découverte [citas de la tr. esp. *El método en sociología*. Córdoba: Ferreyra 2005].
CONVERSE, Jean M. (1984) *Strong Arguments and Weak Evidence: The Open/Closed Questioning Controversy of the 1940s*, en "Public Opinion Quarterly" XLVIII, 1 (spring): 267-82.
CONVERSE, Jean M. y Stanley PRESSER (1986) *Survey Questions: Handcrafting the Standardized Questionnaire*. Beverly Hills: Sage.
CONVERSE, Philip E. (1970) *Attitudes and Non Attitudes: Continuation of a Dialogue*, pp. 168-89 en Edward R. Tufte (comp.), *The Quantitative Analysis of Social Problems*. Reading: Addison-Wesley.
COOK, Stuart W. y Claire SELLTIZ (1964) *A Multiple-Indicator Approach to Attitude Measurement*, en "Psychological Bullettin" LXII, 4 (july): 36-55.
COOK, Thomas D. y Donald T. CAMPBELL (1979) *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Chicago: Rand-McNally.
COOMBS, Clyde H. (1953) *Theory and Methods of Social Measurement*, pp. 471-535 en Leon Festinger y Daniel Katz (comps.), *Research Methods in the Behavioral Sciences*. New York: Dryden.
CORBETTA, Piergiorgio (2003) *Metodología y técnicas de investigación social*. Madrid: Mac Graw-Hill.

CRESPI, Franco (1985) *Le vie della sociologia. Problemi, teorie, metodi*. Bologna: Il Mulino.
CRONBACH, Lee J. (1946) *Response Sets and Test Validity*, en "Educational and Psychological Measurement" VI: 475-94.
CRONBACH, Lee J. (1947) *Test Reliability: Its Meaning and Determination*, en "Psychometrika" XII, 1 (march): 1-16.
CRONBACH, Lee J. (1949) *Essentials of Psychological Testing*. New York: Harper & Row.
CRONBACH, Lee J. (1951) *Coefficient Alpha and the Internal Structure of Tests*, en "Psychometrika" XVI: 297-334.
CRONBACH, Lee J. y Paul E. MEEHL (1955) *Construct Validity in Psychological Tests*, en "Psychological Bulletin" LII, 4 (july): 281-302.
CRONBACH, Lee J. *et al.* (1972) *The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles*. New York: Wiley.
DALTON, Melville E. (1964) *Preconceptions and Methods in "Men Who Manage"*, pp. 58-110 en Phillip E. Hammond (comp.), *Sociologists at Work*. New York: Basic Books.
DAVIES, Alan (1977) *The Construction of Language Tests*, pp. 38-104 en J. P. B. Allen y A. Davies (comps.), *Testing and Experimental Methods*. London: Oxford University Press.
DAVIS, James A. (1964) *'Great Books and Small Groups': An Informal History of a National Survey*, pp. 244-269 en Phillip E. Hammond (comp.), *Sociologists at Work*. New York: Basic Books.
DAVIS, James A. (1971) *Elementary Survey Analysis*. Englewood Cliffs: Prentice-Hall.
DAVIS, James A. (1987) *The Future Study of Public Opinion*, en "Public Opinion Quarterly" LI, 4 (winter): 178-9.
DEI, H. Daniel (2006) *La tesis*. Buenos Aires: Prometeo.
De MAIO, T. J. (1984) *Social Desirability and Survey Measurement : A Review*, pp. 257-82 en C. F. Turner y E. Martin (comps.), *Surveying Subjective Phenomena*. New York: Russell Sage Foundation.
DE MOIVRE, Abraham (1733) *La doctrine des probabilités*. Paris.
DENSTADLI, Jon Martin (2000) *Analyzing Air Travel: A Comparison of Different Survey Methods and Data Collection Procedures*, en "Journal of Travel Research" XXXIX, 1 (agosto).
DENZIN, Norman K. (1970) *The Research Act: A Theoretical Introduction to Sociological Methods*. Chicago: Aldine-Atherton.
DENZIN, Norman K. e Yvona S. LINCOLN (eds.) *Handbook of Qualitative Research*. Thousand Oaks: Sage.
DESCARTES, René (1628) *Regulae ad directionem ingenii*.
DESCARTES, René (1637) *Discours de la Méthode*. Leyden: Ian Maire.
DEWEY, John (1938) *Logic, the Theory of Inquiry*. New York: Holt.
DI FRANCO, Giovanni (2001) *EDS: Esplorare, descrivere, sintetizzare i dati*. Milano: Angeli.
DIJKSTRA, Will y Johannes VAN DER ZOUWEN (comps., 1982) *Response Behavior in the Survey-Interview*. London: Academic Press.
DILLMAN, Don A. (1978) *Mail and Telephone Surveys: The Total Design Method*. New York: Wiley.

DILLMAN, Don A. (2000) *Mail and Internet Surveys: The Tailored Design Method.* New York: Wiley.
DOREIAN, Patrick (1972) *Multivariate Analysis and Categorized Data*, en "Quality & Quantity" VI, 2 (december): 253-72.
DOVRING, Karin (1954-1955) *Quantitative Semantics in 18th Century Sweden*, en "The Public Opinion Quarterly" XVIII, 4: 389-394.
DROYSEN, Johann G. (1868) *Grundriß der Historik.* Leipzig: Veit.
DUNCAN, Otis Dudley (1984) *Notes on Social Measurement. Historical and Critical.* New York: Russell Sage Foundation.
DURKHEIM, Emile (1895) *Les règles de la méthode sociologique.* Paris: Alcan. [Tr. esp. *Las reglas del método sociológico.* Buenos Aires: Hyspamerica 1982].
DURKHEIM, Emile (1896) *Le suicide: Etude de Sociologie.* Paris: Alcan.
DUSTER, Troy (1981) *Intermediate Steps Between Micro- and Macro-Integration: the Case of Screening for Inherited Disorders*, pp. 109-35 en Knorr-Cetina y Cicourel (1981).
EBBINGHAUS, Hermann (1885) *Ueber das Gedaechtnis.* Leipzig: Duncker & Humblot.
ECO, Umberto (1989) *Il pendolo di Foucault.* Milano: Gruppo Editoriale Fabbri, Bompiani, Sonzogno, Etas [tr. esp. *El péndulo de Foucault.* Buenos Aires: Bompiani - Lumen - De la Flor 1989].
EDGERTON, Harold A. y Laverne E. KOLBE (1936) *The Method of Minimum Variation for the Combination of Criteria*, en "Psychometrika" I, 3 (september): 183-77.
EDWARDS, Allen L. (1954) *Personal Preference Schedule.* New York: Psychological Corporation.
EDWARDS, Allen L. (1957) *Techniques of Attitude Scale Construction.* New York: Appleton.
EDWARDS, Allen L. (1957) *The Social Desirability Variable in Personality Assessment and Research.* New York: Dryden.
EDWARDS, Allen L. y Kathryn Claire KENNEY (1946) *A Comparison of the Thurstone and Likert Techniques of Attitude Scale Construction*, en "Journal of Applied Psychology" XXX, 1: 72-83.
FAIRCLOUGH, Norman (1989) *Language and Power.* London: Longman.
FAIRCLOUGH, Norman (1995) *Critical Discourse Analysis.* Harlow: Longman.
FAIRCLOUGH, Norman (2003) *Analysing Discourse: Textual Analysis for Social Research.* London: Routledge.
FECHNER, Gustav (1860) *Elemente der Pshychophysik.*
FERGUSON, A. *et al.* (1935) *Quantitative Estimates of Sensory Events.* Report n. 108, British Association for the Advancement of Science, 331-49.
FERRAND, D. J. y J. M. MARTEL (1986) *Le choix multicritère des items d'une échelle de mesure*, en "Mathématique et Sciences Humaines" n. 89: 35-59.
FERRAROTTI, Franco (1983) *Prefazione*, en Francesco Citarrella, *La sociologia come scienza del singolare. Note sul metodo biografico.* Palermo: Mazzone.
FERRAROTTI, Franco (1986) *La storia e il quotidiano.* Bari: Laterza. [citas de la tr. esp. *La historia y lo cotidiano.* Barcelona: Peninsula 1991].
FEYERABEND, Paul K. (1970) *Against Method. Outline of an Anarchistic Theory of Knowledge*, en Michael E. Radner y Stephen Winokur, S. (eds.), *Analyses of Theory and Methods of Physics and Psychology.* Minneapolis: University of Minnesota Press.
FIDELI, Roberto y Alberto MARRADI (1996) *Intervista*, pp. 71-82 en "Enciclopedia delle Scienze Sociali". Roma: Istituto della Enciclopedia Italiana, vol. V.
FISKE, Donald Winslow (1971) *Measuring the Concepts of Personality.* Chicago: Aldine.
FLECK, Ludwik (1935) *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einfuehrung in die Lehre vom Denkstil und Denkkollektiv.* Basel: Benno Schwabe.
FLESCH, Rudolf (1948) *A New Readability Yardstick*, en "Journal of Applied Psychology" XXXII: 221-233.
FRASER, Nancy y Linda NICHOLSON (1990) *Crítica social sin filosofía: un encuentro entre el feminismo y el posmodernismo*, en Linda NICHOLSON (ed.), *Feminism/Postmodernism.* New York: Routledge.
FREEMAN, Derek (1983) *Margaret Mead and Samoa: The Making and Unmaking of an Anthropological Myth.* Cambridge: Harvard University Press.
FREY, Frederick W. (1970) *Cross-Cultural Survey Research in Political Science*, pp. 173-294 en Robert T. Holt y John E. Turner (comps.), *The Methodology of Comparative Research.* New York: Free Press.
GALLINO, Luciano (1978) *Dizionario di Sociologia.* Torino: UTET.
GALTON, Francis (1888) *Co-relations and their Measurement, Chiefly from Anthropometric Data*, en "Proceedings of the Royal Society" XLV: 135-40.
GALTUNG, Johan (1966) *Teoría y métodos de la investigación social*, tomo II. Buenos Aires: EUDEBA.
GALTUNG, Johan (1967) *Theory and Methods of Social Research.* London: Allen & Unwin.
GARFINKEL, Harold (1964) *Studies of the Routine Grounds of Everyday Activities*, en "Social Problems" XI: 225-50.
GARFINKEL, Harold (1967) *Studies in Ethnomethodology.* Englewood Cliffs: Prentice-Hall.
GASPERONI, Giancarlo y Francesca GIOVANI (2002) *L'inversione della polarità semantica nelle frasi delle scale Likert*, pp. 123-190 en Marradi y Gasperoni (2002).
GEERTZ, Clifford (1973) *The Interpretation of Cultures.* New York: Basic Books.
GEERTZ, Clifford (1983) *Local Knowledge. Further Essays in Interpretative Anthropology.* New York: Basic Books.
GEERTZ, Clifford (1988) *Works and Lives: The Anthropologist as Author.* Stanford: University Press. [citas de la tr. esp. *El antropólogo como autor.* Buenos Aires: Paidós 1989].
GIAMPAGLIA, Giuseppe (1985) *La stima dell'attendibilità nella costruzione delle scale*, en "Rassegna Italiana di Sociologia" XXVI, 2 (aprile): 169-188.
GIAMPAGLIA, Giuseppe (1986) *Alfa, omega e teta: è attendibile la misura dell'attendibilità?*, en "Sociologia e Ricerca Sociale" VII, 21: 75-99.
GIAMPAGLIA, Giuseppe (1990) *Lo scaling unidimensionale nella ricerca sociale.* Napoli: Liguori.
GIDDENS, Anthony (1979) *Central Problems in Social Theory. Action, Structure and Contradiction in Sociological Analysis.* London: Macmillan.

GILLI, Gian Antonio (1971) *Come si fa ricerca.* Milano: Mondadori.
GLASER, Barney G. (1978) *Theoretical Sensitivity: Advances in the Methodology of Grounded Theory.* Mill Valley: Sociology Press.
GLASER, Barney G. y Anselm L. STRAUSS (1967) *The Discovery of Grounded Theory. Strategies for Qualitative Research.* Chicago: Aldine.
GOBO, Giampietro (1997) *Le risposte e il loro contesto. Processi cognitivi e comunicativi nelle interviste standardizzate.* Milano: Angeli.
GOLD, Raymond L. (1958) *Roles in Sociological Field Observations,* en "Social Forces" XXXVI, 3: 217-223.
GOODENOUGH, Ward H. (1944) *A Technique for Scale Analysis,* en "Educational and Psychological Measurement" IV, 3 (autumn): 179-90.
GORDEN, Rymond (1969) *Interviewing. Strategies, Techniques and Tactics.* Homewood: Dorsey Press.
GREENLEAF, Eric (1992) *Measuring Extreme Response Style,* en "Public Opinion Quarterly" LVI: 328-51.
GUBA, E. G. y Yvona S. LINCOLN (1994) *Competing Paradigms in QualitativeResearch,* en Norman K. Denzin e Yvona S. Lincoln (eds.), *Handbook of Qualitative Research.* Thousand Oaks: Sage.
GUBER, Rosana (1991) *El salvaje metropolitano. Reconstrucción del conocimiento social en el trabajo de campo.* Buenos Aires: Legasa [citas de la primera reimpresión de la edición de Paidós, 2005].
GULLIKSEN, Harold O. (1936) *The Content Reliability of a Test,* en "Psychometrika" I, 3 (september): 189-94.
GUTTMAN, Louis A. (1944) *A Basis for Scaling Qualitative Data,* en "American Sociological Review" IX: 139-50.
GUTTMAN, Louis A. (1950) *The Basis for Scalogram Analysis,* pp. 60-90 en Samuel Stouffer (comp.), *Measurement and Prediction,* vol. IV. Princeton University Press.
HABERMAS, Jürgen (1962) *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society.* Cambridge: Polity 1989.
HAMILTON, D. (1994) *Traditions, Preferences, and Postures in Applied Qualitative Research,* en Norman K. Denzin e Yvona S. Lincoln (eds.), *Handbook of Qualitative Research.* Thousand Oaks: Sage.
HAMMERSLEY, Martyn (1989) *The Dilemma of Qualitative Method. Herbert Blumer and the Chicago Tradition.* London: Routledge.
HEISE, David R. y George W. BOHRNSTEDT (1970) *Validity, Invalidity, and Reliability,* pp. 104-129 en Edgar F. Borgatta y George W. Bohrnstedt (comps.), *Sociological Methodology 1970.* San Francisco: Jossey-Brass.
HEMPEL, Carl Gustav (1935) *L'analyse logique de la psychologie,* en "Revue de Synthèse" [citas de la tr. ingl. *The Logical Analysis of Psychology,* pp. 373-84 en Herbert Feigl y Wilfred Sellars (comps.), *Readings in Philosophical Analysis.* New York: Appleton 1949].
HEMPEL, Carl Gustav (1952) *Fundamentals of Concept Formation in Empirical Science.* Chicago University Press.
HEMPEL, Carl Gustav (1966) *Philosophy of Natural Science.* Englewood Cliffs: Prentice Hall.
HEMPEL, Carl Gustav y Paul OPPENHEIM (1936) *Der Typusbegriff im Lichte der neuen Logik.* Leyden.

HOMAN, Roger (1978) *Interpersonal Communication in Pentecostal Meetings,* en "Sociological Review" XXVI, 3: 499-518.
HOMAN, Roger (1980) *The Ethics of Covert Methods,* en "British Journal of Sociology" XXXI, 1: 46-59.
HOMANS, George Caspar (1970) *Bringing Men Back,* en D. Forcese y S. Richer (comps.), *Stages of Social Research.* Englewood Cliffs: Prentice Hall.
HUI, Harry C. y Harry C. TRIANDIS (1985) *The Instability of Response Sets,* en "Public Opinion Quarterly" XLIX, 2 (summer): 253-60.
HYMAN, Herbert H. (1955) *Survey Design and Analysis.* Glencoe: The Free Press.
HYMAN, Herbert H. (1972) *Secondary Analysis of Sample Surveys. Principles, Procedures and Potentialities.* New York: Wiley.
HYMAN, Herbert H. y Paul B. SHEATSLEY (1954) *The Authoritarian Personality: A Methodological Critique,* pp. 50-122 en Richard Christie y Marie Jahoda (comps.), *Studies in the Scope and Method of "The Authoritarian Personality".* Glencoe: Free Press.
INGRAM, Elisabeth (1977) *Basic Concepts in Testing,* pp. 11-37 en J. P. B. Allen y Alan Davies (comps.), *Testing and Experimental Methods.* London: Oxford University Press.
JAMOUS, Haroun (1968) *Technique, méthode, épistémologie,* en "Epistémologie Sociologique", 6 (2ᵉ semestre): 21-38.
JEVONS, William Stanley (1874) *The Principles of Science. A Treatise on Logic and Scientific Method.* London: MacMillan.
JICK, Todd D. (1979) *Mixing Qualitative and Quantitative Methods: Triangulation in Action,* en "Administrative Science Quarterly" XXIV, 4: 602-611.
JOHNSON, T. *et al.* (2005) *The Relation Between Culture and Response Styles: Evidence From 19 Countries,* en "Journal of Cross-Cultural Psychology" XXXVI, 2 (March): 264 –277.
JONES, Steven G. (*ed.,* 1995) *Cybersociety: Computer Mediated Communication and Community.* London: Sage.
JONES, Steven G. (1997) *Virtual Culture Identity and Communication in Cyberspace.* London: Sage.
JORDAN, Nehemiah (1965) *The Asymmetry of "Liking" and "Disliking": a Phenomenon Meriting Further Reflection and Research,* en "Public Opinion Quarterly" XXIX, 2 (summer): 315-22.
JORGENSEN, Danny L. (1989) *Participant Observation: A Methodology for Human Studies.* London: Sage.
JUNG, Carl Gustav (1910) *The Association Method,* en "American Journal of Psychology" XXI: 219-69.
KANT, Immanuel (1781) *Kritik der reinen Vernunft.* Königsberg: Reiger.
KAPLAN, Abraham (1964) *The Conduct of Inquiry. Methodology for Behavioral Science.* San Francisco: Chandler.
KATONA, George (1951) *Psychological Analysis of Economic Behavior.* New York: McGraw-Hill.
KAWULICH, Barbara B. (2005) *Participant Observation as a Data Collection Method,* en "Forum Qualitative Sozialforschung / Forum: Qualitative Social Research" VI, 2: art. 43. <http://www.qualitative-research.net/fqs-texte/2-05/05-2-43-e.htm>.

KEMENY, John G. (1959) *A Philosopher Looks at Science.* Princeton: Van Nostrand [citas de la tr. it. *Il filosofo e la scienza. Introduzione alla filosofia della scienza.* Milano: Il Saggiatore 1972].
KENDALL Maurice G. (1948) *Rank Correlation Methods.* London: Griffin.
KINSEY, Alfred Charles *et al.* (1948) *Sexual Behavior in the Human Male.* Philadelphia: Sounders.
KINSEY, Alfred Charles *et al.* (1953) *Sexual Behavior in the Human Female.* Philadelphia: Sounders.
KITZINGER, Jenny (1995) *Introducing Focus Groups* en "British Medical Journal" n. 311: 299-302.
KLIMOVSKY, Gregorio (1994) *Las desventuras del conocimiento científico. Una introducción a la epistemología.* Buenos Aires: AZ Editora.
KNAFL, Katheleen A. y Bonnie J. BREITMAYER (1989) *Triangulation in Qualitative Research: Issues of Conceptual Clarity and Purpose*, pp. 209-220 en Janice M. Morse (*ed.*), *Qualitative Nursing Research: A Contemporary Dialogue.* Rockville: Aspen.
KNORR-CETINA, Karin D. y Aaron Victor CICOUREL (eds., 1981) *Advances in Social Theory and Methodology. Toward an Integration of Micro- and Macro-Sociologies.* London: Routledge & Kegan Paul.
KRIPPENDORFF, Klaus (1980) *Content Analysis. An Introduction to its Methodology.* Newbury Park: Sage.
KRIZ, Juergen (1988) *Facts and Artifacts in Social Science. An Epistemological and Methodological Analysis of Empirical Social Science Research Techniques.* New York: MacGraw-Hill.
KREUGER, Richard A. (1988) *Focus Groups: A Practical Guide for Applied Research.* London: Sage.
KUDER, G. Frederick y Marion W. RICHARDSON (1937) *The Theory of the Estimation of Test Reliability*, en "Psychometrika" II, 3 (september): 151-60.
KUHN, Thomas Samuel (1962) *The Structure of Scientific Revolutions.* Chicago University Press [citas de la tr. esp. *La estructura de las revoluciones científicas.* Santiago de Chile: FCE 1993].
LABOVITZ, Sanford (1968) *Criteria for Selecting a Significance Level: A Note on the Sacredness of .05*, en "American Sociologist" III, 3 (august): 220-22.
LAU, Richard (1982) *Negativity in Political Perception*, en "Political Behavior" IV: 353-377.
LAZARSFELD, Paul Felix (1937) *Some Remarks on the Typological Procedures in Social Research*, en "Zeitschrift für Sozialforschung" VI: 119-39.
LAZARSFELD, Paul Felix (1955) *Interpretation of Statiscal Relations as a Research Operation*, en Paul Felix Lazarsfeld y Morris Rosenberg (comps.), *The Language of Social Research.* New York: Free Press.
LAZARSFELD, Paul Felix (1958) *Evidence and Inference in Social Research*, en "Daedalus" LXXXVII, 3 (fall): 99-130.
LAZARSFELD, Paul Felix (1959) *Problems in Methodology*, pp. 39-78 en Robert King Merton, L. Broom y L. S. Cottrell (comps.), *Sociology Today: Problems and Prospects.* New York: Basic Books [citas de la tr. it. *Problemi di metodologia*, pp. 179-229 en Lazarsfeld, *Metodologia e ricerca sociologica.* Bologna: Il Mulino 1967].

LAZARSFELD, Paul Felix (1966) *Concept Formation and Measurement in the Behavioral Sciences: Some Historical Observations*, en Gordon J. Direnzo (comp.), *Concepts, Theories and Explanation in the Behavioral Sciences.* New York: Random.
LAZARSFELD, Paul Felix (1972) *Preface*, en Lazarsfeld *et al.* (comps.), *Continuities in the Language of Social Research.* New York: Free Press.
LAZARSFELD, Paul Felix y Allen H. BARTON (1951) *Qualitative Measurement in the Social Sciences: Classifications, Typologies, and Indices*, pp. 155-92 en Daniel Lerner y Harold D. Lasswell (comps.), *The Policy Sciences.* Stanford University Press.
LAZARSFELD, Paul Felix, Bernard BERELSON y Hazel GAUDET (1944) *The People's Choice. How the Voter Makes up His Mind in a Presidential Campaign.* New York: Duell Sloan & Pearce.
LAZARSFELD, Paul Felix, Bernard BERELSON y William MCPHEE (1954) *Voting: A Study in a Presidential Campaign.* University of Chicago Press.
LAZARSFELD, Paul Felix y Morris ROSENBERG (1955) *Introduction*, pp. 1-12 en Lazarsfeld y Rosenberg (comps.), *The Language of Social Research.* New York: Free Press.
LEBART, Ludovic y André SALEM (1994) *Statistique Textuelle.* París: Dunod.
LECUYER, Bernard-Pierre (1968) *Histoire et sociologie de la recherche sociale empirique: problèmes de théorie et de méthode*, en "Epistémologie Sociologique" n. 6 (28 sémestre): 119-31.
LEONARDI, Franco (1991) *Contro l'analisi qualitativa*, en "Sociologia e ricerca sociale" 35: 3-29.
LEROY, Edouard (1899) *Science et philosophie*, en "Revue de métaphysique et morale" VII, 2e partie: 364-381.
LIKERT, Rensis (1932) *The Method of Constructing an Attitude Scale*, pp. 44-53 en Rensis Likert, *A Technique for the Measurement of Attitudes*, en "Archives of Psychology", monogr. n. 140.
LIPSET, Seymour Martin (1959) *Some Social Requisites of Democracy: Economic Development and Political Legitimacy*, en "American Political Science Review" LIII, 69-105.
LORD, Frederick M. y Melvin R. NOVICK (1968) *Statistical Theory of Mental Test Scores.* Reading: Addison-Wesley.
LOSEE, John (1972) *A Historical Introduction to the Philosophy of Science.* Oxford University Press.
LOSITO, Gianni (1993) *L'analisi del contenuto nella ricerca sociale.* Milano: Angeli.
LULO, Jorge (2002) *La vía hermenéutica: las ciencias sociales entre la epistemología y la ontología*, pp. 177-235 en Federico L. Schuster (comp.), *Filosofía y métodos de las ciencias sociales.* Buenos Aires: Manantial.
LUNDBERG, George A. (1938) *The Concept of Law in the Social Sciences*, en "Philosophy of Science" V, 2 (april): 189-203.
LUTYNSKI, Jan (1978) *Koncepcje pytania kwestionariuszowego, ich zastowanie w badaniach, konsekwencje i mozliwosci*, en "Przeglad Socjologiczny" XXX: 9-51.
MADGE, John (1962) *The Origins of Scientific Sociology.* New York: Free Press [citas de la tr. it. *Lo sviluppo dei metodi di ricerca empirica in sociologia.* Bologna: Il Mulino 1966].

MALTHUS, Thomas Robert (1747) *An Essay on the Principles of Population, as it Affects the Future Improvement of Society.* London: Johnson.

MARRADI, Alberto (1994) *Referenti, pensiero e linguaggio: una questione rilevante per gli indicatori*, en "Sociologia e Ricerca Sociale" XV, 43 (marzo): 137-207.

MARRADI, Alberto (1997) *Casuale e rappresentativo: ma cosa vuol dire?*, pp. 23-87 en Paolo Ceri (comp.), *Politica e sondaggi.* Torino: Rosenberg & Sellier.

MARRADI, Alberto (1997b) *Esperimento, associazione, insieme non standard?* en Giovanni Bettin (comp.), *Politica e Società.* Padova: Cedam.

MARRADI, Alberto (2000) *Método como arte.* "Revista Argentina de Economía y Ciencias Sociales" IV (primavera 2000): 7-25.

MARRADI, Alberto (2002) *Le scale Likert e la reazione all'oggetto*, pp. 15-52 en Marradi y Gasperoni (2002).

MARRADI, Alberto (2005) *Raccontar storie.* Roma: Carocci.

MARRADI, Alberto y Giancarlo GASPERONI (eds., 2002) *Costruire il dato 3. Le scale Likert.* Milano: Angeli.

MARSH, Catherine (1982) *The Survey Method: The Contribution of Surveys to Sociological Explanation.* London: Allen & Unwin.

MARTIRE, Fabrizio (2006) *Come nasce e come cresce una scuola sociologica.* Roma: Bonanno.

MASSEY, Alexander (1999) *Methodological Triangulation, Or How To Get Lost Without Being Found Out*, pp. 183-197 en Alexander Massey y Geoffrey Walford (*eds.*), *Explorations in Methodology, Studies in Educational Ethnography.* Stanford: JAI Press.

MAXWELL, Joseph (1996) *Qualitative Research Design. An Interactive Approach.* Thousand Oaks: Sage.

McKENNELL, Aubrey C. (1970) *Attitude Measurement: Use of Coefficient Alpha with Cluster or Factor Analysis*, en "Sociology" IV, 2 (may): 227-45.

McKENNELL, Aubrey C. (1973) *Surveying Attitude Structures: A Discussion of Principles and Procedures*, en "Quality and Quantity" VII, 2 (december): 203-94.

McKENNELL, Aubrey C. (1977) *Attitude Scale Construction*, pp. 183-219 en Colm O' Muircheartaigh y Clive Payne (comps.), *The Analysis of Survey Data*, vol. I. New York: Wiley.

McNEMAR, Quinn (1946) *Opinion-Attitude Methodology*, en "Psychological Bulletin" XLIII, 4 (july): 289-374.

MEAD, Margaret (1928) *Coming of Age in Samoa.* New York: Morrow.

MERRITT, Richard L. y Stein ROKKAN (comps., 1966) *Comparing Nations. The Use of Quantitative Data in Cross-National Research.* New Haven: Yale University Press.

MERTON, Robert King (1940) *Fact and Factitiousness in Ethnic Opinionaires*, en "American Sociological Review" V, 1: 13-27.

MERTON, Robert King (1948) *The Bearing of Empirical Research Upon the Development of Social Theory*, en "American Sociological Review" XIII, 5 (october): 505-15.

MERTON, Robert King (1949) *Social Theory and Social Structure. Toward the Codification of Theory and Research.* Glencoe: Free Press [algunas citas de la traducción *Teoría y estructura sociales.* México: FCE 1964].

MERTON Robert King y Patricia L. KENDALL (1946) *The Focused Interview*, en "American Journal of Sociology" LI, 6: 541-557.

MILLER, George Armitage (1956) *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*, en "Psychological Review" LXIII: 81-97.

MILLER, Kevin W. y Lora B. WILDER (1997) *The Feasibility of a Street-Intercept Survey Method in an African-American Community*, en "American Journal of Public Health" LXXXVII, 4 (april): 655.

MILLS, Charles Wright (1959) *The Sociological Imagination.* New York: Oxford University Press.

MISHLER, Elliot G. (1986) *Research Interviewing.* Cambridge University Press.

MOKKEN, Robert (1971) *A Theory and Procedure of Scale Analysis.* Paris: Mouton.

MOKRZYCKI, Edmund (1983) *Philosophy of Science and Sociology. From the Methodological Doctrine to Research Practice.* London: Routledge & Kegan Paul.

MONTESPERELLI, Paolo (1998) *L'intervista ermeneutica.* Milano: Angeli.

MONTESPERELLI, Paolo (2005) *Sociología de la memoria.* Buenos Aires: Ediciones Nueva Visión.

MORGAN, David L. (1988) *Focus Groups as Qualitative Research.* London: Sage.

MORGAN, David L. y Richard A. KREUGER (1993) *When to Use Focus Groups and Why*, en David L. Morgan (comp.), *Successful Focus Groups: Advancing the State of the Art.* London: Sage.

MULKAY, Michael J. (1979) *Science and the Sociology of Knowledge.* London: Allen & Unwin.

MURDOCK, George Peter (1949) *Social Structure.* New York: MacMillan.

MURRAY, Henry A. (1938) *Explorations in Personality.* New York: Oxford University Press.

NADEL, Siegfried Frederick (1949) *The Foundations of Social Anthropology.* London: Cohen & West [citas de la tr. it. *Lineamenti di antropologia sociale.* Bari: Laterza 1974].

NAGEL, Ernst (1961) *The Structure of Science: Problems in the Logic of Scientific Explanation.* London: Routledge & Kegan Paul.

NAROLL, Raoul (1968) *Some Thoughts on Comparative Method in Cultural Anthropology*, pp. 236-277 en Hubert M. Blalock y Ann B. Blalock (comps.), *Methodology in Social Research.* New York: McGraw-Hill.

NEURATH, Otto (1931-32) *Soziologie im Physikalismus*, en "Erkenntnis" II: 393-431.

NOWAK, Stefan (1976) *Understanding and Prediction. Essays in the Methodology of Social and Behavioral Theories.* Dordrecht: Reidel.

ORVIK, J. M. (1972) *Social Desirability for the Individual, his Group and Society*, en "Multivaríate Behavioral Research" VII: 3-32.

OSGOOD, Charles E. (1952) *The Nature and Measurement of Meaning*, en "Psychological Bulletin" XLIX, 3 (may): 197-237.

OSGOOD, Charles E. *et al.* (1957) *The Measurement of Meaning.* Urbana: University of Illinois Press.

OSORIO, Francisco *et al.* (1999) *Observación Participante. El caso de Alison Spedding*, en "Cinta de Moebio" n. 6. <http://rehue.csociales.uchile.cl/publicaciones/moebio/06/spedding.htm>.

OUTHWAITE, William (1987) *New Philosophies of Social Science.* Basingstoke: Macmillan.

PALUMBO, Mauro y Elisabetta GARBARINO (2004) *Strumenti e strategie della ricerca sociale.* Milano: Franco Angeli.

PARISI, Domenico y Cristiano CASTELFRANCHI (1978) *Una definizione della psicologia cognitivista,* pp. 63-84 en Gaetano Kanizsa y Paolo Legrenzi (eds.), *Psicologia della gestalt e psicologia cognitivista.* Bologna: Il Mulino.

PARRY, Hugh y Helen M. CROSSLY (1950) *Validity of Responses to Survey Questions,* en "Public Opinion Quarterly" XIV, 1 (spring): 61-80.

PARSONS, Talcott (1937) *The Structure of Social Action. A Study in Sociological Theory with Special Reference to a Group of Recent European Writers.* New York: McGraw-Hill.

PATTON, Michael Q. (1986) *Utilization-focused Evaluation.* London: Sage.

PAVSIC, Rita y Maria Concetta PITRONE (2000) *Conviene rilevare gli atteggiamenti con la* forced choice?, en "Sociologia e Ricerca Sociale" XXI, 62: 81-124.

PAVSIC, Rita y Maria Concetta PITRONE (2003) *Come conoscere opinioni e atteggiamenti.* Roma: Bonanno.

PAWSON, Ray y N. TILLEY (1997) *Realistic Evaluation.* London: Sage.

PAYNE, Stanley L. (1949) *The Art of Asking Questions.* Princeton University Press.

PEABODY, Dean (1961) *Attitude Content and Agreement Set in Scales of Authoritarianism, Dogmatism, Anti-Semitism and Economic Conservatism,* en "Journal of Abnormal and Social Psychology" LXIII, 1: 1-18.

PEARSON, Karl (1896) *Mathematical Contributions to the Theory of Evolution: Regression, Heredity, and Panmixia,* en "Philosophical Transactions of the Royal Society of London" CXXXVII: 253-318.

PEARSON, Karl (1913) *On the Measurement of the Influence of "Broad Categories" on Correlation,* en "Biometrika" IX: 116-139.

PERA, Marcello (1991) *Scienza e retorica.* Bari: Laterza.

PERA, Marcello (1978) *Induzione e metodo scientifico.* Pisa: Editrice Tecnico-Scientifica.

PEREZ RANSANZ, Ana Rosa (1995) *Racionalidad y desarrollo científico,* pp. 171-201 en León Olive (comp.), *Racionalidad epistémica.* Madrid: Trotta.

PILICHOWSKI, Andrzej y Wlodzimierz ROSTOCKI (1978) *Powtorny wywiad werifikacyiny jako metoda otrzymywania informacji o wartosci opdowiedzi na pytanie kwestionariuszowe,* en "Przeglad Socjologiczny" XXX: 69-78.

PIOVANI, Juan Ignacio (2002) *Idee di scienza nella comunità accademica inglese,* en "Sociologia e Ricerca Sociale" XXIII, 37: 91-116.

PIOVANI, Juan Ignacio (2006) *Alle origini della statistica moderna. La scuola inglese di fine Ottocento.* Milano: Angeli.

PIOVANI, Juan Ignacio (coord.) *et al.* (2006) *Producción y reproducción de sentidos en torno a lo cualitativo y lo cuantitativo en la sociología,* "I Foro de Metodologías y Prácticas de Investigación Social", La Plata y Buenos Aires: UNLP/ UBA.

PLATT, Jennifer (1983) *The Development of the "Participant Observation" Method in Sociology: Origin Myth and History,* en "Journal of the History of the Behavioral Sciences" XIX, 4: 379-393.

PLATT, Jennifer (1985) *Weber's Verstehen and the History of Qualitative Research: The Missing Link,* en "The British Journal of Sociology" XXXVI, 3: 448-466.

PLUMMER, Ken (1983) *Documents of Life: An Introduction to the Problems and Literature of a Humanistic Method.* London: Allen & Unwin.

POLANYI, Michael (1958) *Personal Knowledge. Towards Post-Critical Philosophy.* London: Routledge & Kegan Paul.

POPPER, Karl Raimund (1934) *Logik der Forschung.* Wien: Springer.

POPPER, Karl Raimund (1963) *Conjectures and refutations.* London: Routledge.

POPPER, Karl Raimund (1972) *Objective Knowledge. An Evolutionary Approach.* Oxford: Clarendon Press.

POWELL, Richard A. y Helen M. SINGLE (1996) *Focus groups,* en "International Journal of Quality in Health Care" VIII, 5: 499-504.

PRESSER, Stanley y Michael TRAUGOTT (1992) *Little White Lies and Social Science Models,* en "Public Opinion Quarterly" LVI, 1: 77-86.

PRZEWORSKI, Adam y Henry TEUNE (1970) *The Logic of Comparative Social Inquiry.* New York: Wiley.

RASCH, George (1980) *Probabilistic Models for Some Intelligence and Attainment Tests.* Chicago University Press.

RAVETZ, Jerome R. (1971) *Scientific Knowledge and its Social Problems.* Oxford: Clarendon Press.

RAY, John J. (1990) *Acquiescence And Problems With Forced-Choice Scales,* en "Journal of Social Psychology" CXXX, 3: 397-399.

REBELLO, Lorraine M. (1977) *Observation of Human Behavior. A Methodological Review,* en "The Indian Journal of Social Work" XXXVII, 4: 357-371.

RICOLFI, Luca (1982) *Modelli di analisi e software per le scienze sociali,* en "Rassegna Italiana di Sociologia" XXIII, 3 (luglio): 331-72.

ROBINSON, John P. *et al.* (comps., 1991) *Measures of Social Psychological Attitudes.* San Diego: Academic Press.

ROETHLISBERGER, Fritz Jules y William John DIKSON (1939) *Management and the Worker. Technical vs. Social Organization in an Industrial Plant.* Cambridge: Harvard University Press.

ROGERS, Carl R. (1942) *Counselling and Psychotherapy.* Boston: Houghton Mifflin.

RORSCHACH, Hermann (1921) *Psychodiagnostic: A Diagnostic Test Based on Perception.* New York: Grune & Stratton.

ROSENBLUM, Karen E. (1987) *The In-Depth Interview: Between Science and Sociability,* en "Social Forces" II, 2: 388-400.

ROSKAM, Edward E. (1989) *Operationalization, a Superfluous Concept,* en "Quality & Quantity" XXIII, 3-4: 237-75.

ROSENZWEIG, Saul (1938) *The Experimental Measurement of Types of Reaction to Frustration,* pp. 585-99 en H. A. Murray *et al.* (comps.), *Explorations in Personality.* New York: Oxford University Press.

ROTH, Julius (1965) *Hired Hand Research,* en "American Sociologist" I, 1 (november): 190-6.

RUDNER, Richard S. (1966) *Philosophy of Social Science.* Englewood Cliffs: Prentice-Hall [citas de la tr. it. *Filosofia delle scienze sociali.* Bologna: Il Mulino 1968].

RUGG, Donald y Hadley CANTRIL (1944) *The Wording of Questions*, pp. 23-50 en Hadley Cantril (comp.), *Gauging Public Opinion*. Princeton University Press.
RUIZ OLABUENAGA, José I. (1996) *Metodología de la investigación cualitativa*. Bilbao: Universidad de Deusto.
RUIZ OLABUENAGA, José I. y María A. ISPIZUA (1989) *La descodificación de la vida cotidiana. Métodos de investigación cualitativa*. Bilbao: Universidad de Deusto.
RUMMEL, Rudolph J. (1963) *Dimensions of Conflict: Behavior Within and Between Nations*, en "General Systems Yearbook" VIII: 1-50.
SAMAJA, Juan (1999) *Epistemología y metodología. Elementos para una teoría de la investigación científica*. Buenos Aires: EUDEBA.
SAMAJA, Juan (2002a) *Análisis del proceso de investigación*, pp. 199-235 en H. Daniel Dei (ed.), *Pensar y hacer en investigación*. Buenos Aires: Docencia.
SAMAJA, Juan (2002b) *Concepto, método y sentido de la investigación científica*, pp. 157-180 en H. Daniel Dei (ed.), *Pensar y hacer en investigación*. Buenos Aires: Docencia.
SANDELOWSKI, Margarete y Julie BARROSO (2002) *Reading Qualitative Studies*, en "International Journal of Qualitative Methods" I, 1: art. 5.
SAPIGNOLI, Michele (2002) *Ricerche sulla reazione all'oggetto a altri tipi di distorsione*, pp. 53-93 en Marradi y Gasperoni (2002).
SARTORI, Giovanni (1984) *Foreword* y *Guidelines for Concept Analysis*, pp. 1-85 en Giovanni Sartori (comp.), *Social Science Concepts. A Systematic Analysis*. London: Sage.
SAUTU, Ruth *et al.* (2005) *Manual de metodología*. Buenos Aires: CLACSO.
SAVIANE, Sergio (1978) *Queste povere puttane*, en "L'Espresso" 17.9.
SCHATZMAN, Leonard y Anselm STRAUSS (1973) *Field Research: Strategies for a Natural Science*. Englewood Cliffs: Prentice-Hall.
SCHUMAN, Howard (1966) *The Random Probe. A Technique for Evaluating the Validity of Closed Questions*, en "American Sociological Review" XXV, 1 (february): 3-25.
SCHUMAN, Howard y Stanley PRESSER (1981) *Questions and Answers in Attitude Surveys. Experiments on Question Form, Wording, and Context*. New York: Academic Press.
SCHUMPETER, Joseph A. (1942) *Capitalism, Socialism, and Democracy*. New York: Harper & Row.
SCHUSSLER, Karl (1982) *Measuring Social Life Feelings*. San Francisco: Jossey-Bass.
SCHUTZ, Alfred (1945) *On Multiple Realities*, en "Philosophy and Phenomenological Research", V (june): 533-75.
SCHUTZ, Alfred (1953) *Common Sense and Scientific Interpretation of Human Action*, en "Philosophy and Phenomenological Research" XIV, 1 (september): 1-37.
SCHUTZ, Alfred (1962) *The Problem of Social Reality*, editado por Maurice Natanson. The Hague: Nijoff.
SCHWANDT, Thomas (1994) *Constructivist, Interpretivist Approaches to Human Inquiry*, pp. 118-137 en Norman K. Denzin y Y. S. Lincoln (comps.), *Handbook of Qualitative Research*. Thousand Oaks: Sage.

SCOTT, William A. (1968) *Attitude Measurement*, pp. 204-273 en Gardner Lindzey y Elliot Aronson (comps.), *Handbook of Social Psychology*, vol. II. Reading: Addison-Wesley.
SEBASTIANI, Chiara (1970) *Marginalità politica e integrazione manipolata: sondaggio in tre borgate romane*, en "La Critica Sociologica" n. 14 (estate): 89-134.
SELVIN, Hanan C. (1958) *Durkheim's Suicide and Problems of Empirical Research*, en "American Journal of Sociology" LXIII: 607-19.
SHAW, Marvin E. y Jack Mason WRIGHT (1967) *Scales for the Measurement of Attitudes*. New York: McGraw-Hill.
SINGER, J. David (1982) *Variables, Indicators, and Data*, en "Social Science History" VI, 2 (spring): 181-217.
SINGLETON, Royce A. Jr. *et al.* (1988) *Approaches to Social Research*. New York: Oxford University Press [citas de la segunda edición, 1993].
SKINNER, Burrhus Frederick (1953) *Science and Human Behaviour*. Glencoe: Free Press.
SMELSER, Neil J. (1976) *Comparative Methods in the Social Sciences*. Englewood Cliffs: Prentice-Hall.
SMITH, A. K. (1969) *Socio-Economic Development and Political Democracy: A Causal Analysis*, en "Midwest Journal of Political Science" XIII, 1 (february): 100-5.
SOLDANO, Daniela (2002) *La subjetividad en escena. El aporte de Alfred Schütz a las ciencias sociales*, pp. 59-92 en Federico L. Schuster (comp.), *Filosofía y métodos de las ciencias sociales*. Buenos Aires: Manantial.
SPEARMAN, Charles Edward (1904) *General Intelligence, Objectively Determined and Measured*, en "American Journal of Psychology" XV: 201-93.
SPEARMAN, Charles Edward (1904b) *The Proof and Measurement of Association Between Two Things*, en "American Journal of Psychology" XV, 1: 72-101.
SPITTLER, Gerd (2001) *Teilnehmende Beobachtung als Dichte Teilnahme*, en "Zeitschrift fur Ethnologie" CXXVI, 1: 1-25.
SPRADLEY, James P. (1980) *Participant Observation*. New York: Holt, Rinehart & Winston.
SPROULL, Lee S. (1986) *Using Electronic Mail for Data Collection in Organizational Research* en "Academy of Management Journal" XXIX, 1: 159-69.
SRINIVAS, Mysore N. (1966) *Social Change in Modern India*. Berkeley: University of California Press.
SRINIVAS, Mysore N. (1998) *Village, Caste, Gender and Method: Essays in Indian Social Anthropology*. Delhi: Oxford University Press.
STATERA, Gianni (1982) *Metodologia e tecniche della ricerca sociale. Una introduzione sistematica*. Palermo: Palumbo.
STATERA, Gianni (1984) *Contro la "Nouvelle vague" antimetodologica*, in "Sociologia e Ricerca Sociale" V, 13 (aprile): 243-61.
STATERA, Gianni (1992) *Il mito della ricerca qualitativa*, en "Sociologia e ricerca sociale" XIII, 39: 5-28.
STEMLER, Steve (2001) *An Overview of Content Analysis*, en "Practical Assessment, Research & Evaluation" VII, 17. <http://pareonline.net/getvn.asp?v=7&n=17>.
STERN, Alan J. (1975) *Rudimentary Political Beliefs in Four Italian Communities*, en "Journal of Politics" XXXVII: 235-60.

STEVENS, Stanley Smith (1946) *On the Theory of Scales of Measurement*, en "Science" CIII, 2684 (june 7): 677-680.
STEVENS, Stanley Smith (1951) *Mathematics, Measurement and Psychophysics*, pp. 1-49 en S. Stevens (comp.), *Handbook of Experimental Psychology*. New York: Wiley.
STEWART, David W. y Prem N. SHAMDASANI (1990) *Focus Groups: Theory and Practice*. London: Sage.
STINCHCOMBE, Arthur (1968) *Constructing Social Theories*. New York: Harcourt Brace & World [citas de la tr. esp. *La construcción de teorías sociales*. Buenos Aires: Nueva Visión 1970].
STOETZEL, Jean (1965) *Préface* a Raymond Boudon y Paul F. Lazarsfeld (comps.), *Méthodes de la sociologie*. Paris: Mouton [citas de la tr. it. *Prefazione all'edizione francese*, pp. 9-14 en Raymond Boudon y Paul F. Lazarsfeld (comps.), *L'analisi empirica nelle scienze sociali*. Bologna: Il Mulino 1969].
STRASSER, Carlos (1979) *La razón científica en política y sociología*. Buenos Aires: Amorrortu.
STRAUSS, Anselm y Juliet CORBIN (1994) *Grounded Theory Methodology: An Overview*, pp. 273-285 en Norman K. Denzin y Y. S. Lincoln (comps.), *Handbook of Qualitative Research*, London: Sage.
STUBBS, Michael (1983) *Discourse Analysis: The Sociolinguistic Analysis of Natural Language*. Chicago: University Press.
SUDMAN, Seymour y Norman M. BRADBURN (1982) *Asking Questions. A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass 1982.
SULLIVAN, John L. (1974) *Multiple Indicators: Some Criteria of Selection*, pp. 243-69 en Hubert M. Blalock (comp.), *Measurement in the Social Sciences*. Chicago: Aldine.
SULLIVAN, John L. y Stanley FELDMAN (1979) *Multiple Indicators. An Introduction*. London: Sage.
SYKES, Wendy (1991) *Taking Stock: Issues From the Literature on Validity and Reliability*, en "Qualitative Research, International Journal of Market Research" XXXIII, 1.
TAYLOR, Charles Lewis y Michael C. HUDSON (1972) *World Handbook of Political and Social Indicators, II*. New Haven: Yale University Press.
TEUNE, Henry (1968) *Measurement in Comparative Research*, en "Comparative Political Studies" I, 1 (april): 123-38.
THEODORSON, George A. y Achilles G. THEODORSON (1970) *A Modern Dictionary of Sociology*. London: Methuen.
THOMAS, William I. y Florian W. ZNANIECKI (1918-1920) *The Polish Peasant in Europe and America*: Chicago: University Press.
THURSTONE, Louis Leon (1927) *The Method of Paired Comparison for Social Values*, en "Journal of Abnormal and Social Psychology" XXI: 384-97.
THURSTONE, Louis Leon (1928) *Attitudes Can be Measured*, en "American Journal of Sociology" XXXII, 4 (January): 529-54.
THURSTONE, Louis Leon (1931) *Rank Order as a Psychophysical Method*, en "Journal of Experimental Psychology" VI: 187-201.
THURSTONE, Louis Leon y Ernest John CHAVE (1929) *The Measurement of Attitude*. University of Chicago Press.

TOLLER, Cinzia (1994) *Memorie di un'intervistatrice pentita*, en "Politica ed economia" XXV, 5-6 (settembre): 85-90.
TONER, B. (1987) *The Impact of Agreement Bias on the Ranking of Questionnaire Response*, en "Journal of Social Psychology" CXXVII: 221-222.
TOULMIN, Stephen Edelston (1972) *Human Understanding*, vol. I: *The Collective Use and Evolution of Concepts*, Princeton University Press.
TUFTE, Edward R. (1970) *Improving Data Analysis in Political Science*, pp. 437-449 en Tufte (comp.), *The Quantitative Analysis of Social Problems*. Reading: Addison-Wesley.
TUKEY, John W. (1954) *Causation, Regression and Path Analysis*. Ames: Iowa State College Press.
TURKLE, Sherry (1995) *Life on the Screen: Identity in the Age of the Internet*. New York: Simon & Schuster.
TURNER, Stephen P. (1979) *The Concept of Face Validity*, en "Quality & Quantity" XIII, 1 (february): 85-90.
VALLES, Miguel S. (1997) *Técnicas cualitativas de investigación social. Reflexión metodológica y práctica profesional*. Madrid: Síntesis.
VAN DIJK, Teun A. (1977) *Text and Context. Explorations in the Semantics and Pragmatics of Discourse*. London: Longman.
VAN DIJK, Teun A. (1978) *Tekstwetenschap. Een interdisciplinaire inleiding*. Utrecht: Het Spectrum.
VAN DIJK, Teun A. (1988) *News as Discourse*. Hillsdale: Erlbaum.
VAN DIJK, Teun A. (1991) *Racism and the Press*. London: Routledge.
VAN DIJK, Teun A. (1998) *Ideology*. London: Sage.
VAN DIJK, Teun A. (2003) *Dominación étnica y racismo discursivo en España y América Latina*. Barcelona: Gedisa.
VAN LEEUWEN, Theo (2005) *Introduction to Social Semiotics*. London: Routledge.
VAN METER, Karl M. (1994) *Sociological Methodology*, en "International Social Science Journal" XLVI, 139 (february): 15-25.
VASILACHIS, Irene (1992) *El análisis lingüístico en la recolección e interpretación de materiales cualitativos*, pp. 153-210 en Floreal Forni *et al.*, *Métodos Cualitativos II. La práctica de la investigación*. Buenos Aires: CEAL.
VASILACHIS, Irene (1997) *Discurso político y prensa escrita*. Buenos Aires: Gedisa.
VASILACHIS, Irene (2000) *Del sujeto cognoscente al sujeto conocido: una propuesta epistemológica y metodológica para el estudio de los pobres y de la pobreza*, 217-245 en Vasilachis (comp.), *Pobres, pobreza y exclusión social*. Buenos Aires: CEIL.
VERBA, Sidney (1969) *The Uses of Survey Research in the Study of Comparative Politics: Issues and Strategies*, pp. 56-105 en Stein Rokkan *et al.* (comps.), *Comparative Survey Analysis*. Paris: Mouton.
VIDICH, Arthur J. y Stanford M. LYMAN (1994) *Qualitative Methods: Their History in Sociology and Anthropology*, en Norman K. Denzin e Yvonna S. Lincoln (eds.), *Handbook of Qualitative Research*. Thousand Oaks: Sage.
WALLIS, W. Allen y Harry V. ROBERTS (1956) *Statistics: A New Approach*. New York: Free Press.
WEBB, Eugene J. *et al.* (1966) *Unobtrusive Measures: Nonreactive Research in the Social Sciences*. Chicago: Rand McNally.

WEBER, Max (1904) *Die Objektivität sozialwissenschaftlicher und sozialpolitischer Erkenntnis*, en "Archiv für Sozialwissenschaft und Sozialpolitik" XIX: 22-87 [citas de la trad. esp. en Weber, *Ensayos sobre metodología sociológica*. Buenos Aires: Amorrortu].
WEBER, Max (1906) *Kritische Studien auf dem Gebiet der Kulturwissenschaftlichen Logik*, en "Archiv für Sozialwissenschaft und Sozialpolitik" XXII: 143-207.
WEBER, Max (1913) *Über einige Kategorien der verstehenden Soziologie*, en "Logos" IV; reproducido en Weber *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen: Mohr 1922.
WEBER, Max (1917) *Der Sinn der 'Wertfreiheit' der soziologischen und ökonomischen Wissenschaften* [citas de la trad. esp. en Weber, *Ensayos sobre metodología sociológica*. Buenos Aires: Amorrortu].
WEBER, Max (1922) *Wirtschaft und Gesellschaft*. Tübingen: Mohr.
WEIGERT, Andrew J. (1970) *The Immoral Rhetoric of Scientific Sociology*, en "American Sociologist" V: 110-23.
WIGGINS, Jerry S. (1968) *Personality Structure*, en "Annual Review of Psychology" XIX: 293-350.
WILCOX, Clyde *et al.* (1989) *Some Like It Hot. Individual Differences in Responses to Group Feeling Thermometers*, en "Public Opinion Quarterly" LIII (summer): 246-57.
WINCH, Peter (1964) *Understanding a Primitive Society*, en "American Philosophical Quarterly" I: 307-24.
WOLIN, Sheldon (1960) *Politics and Vision. Continuity and Innovation in Western Political Thought*. Boston: Little, Brown [citas de la trad. esp. *Política y Perspectiva*. Buenos Aires: Amorrortu 1973].
YULE, George Udny (1900) *On the Association of Attributes in Statistics*, en "Philosophical Transactions of the Royal Society of London" Series A, vol. 194.
ZETTERBERG, Hans L. (1954) *On Theory and Verification in Sociology*. Totowa: Bedminster Press.
ZNANIECKI, Florian W. (1934) *The Method of Sociology*. New York: Farrar & Rinehart.

ÍNDICE ANALÍTICO

Los números indican los capítulos (una cifra) o apartados (dos cifras) en los que se trata un tema